WEIGHT TRANSFER STYLES IN THE GOLF SWING: INDIVIDUAL AND GROUP ANALYSIS

A Research Thesis Presented to

School of Human Movement, Recreation and Performance
Faculty of Human Development

VICTORIA UNIVERSITY

In fulfilment of the requirements for the degree

DOCTOR OF PHILOSOPHY

By

KEVIN BALL

2006

Principal Supervisor: Dr Russell Best
Co-Supervisor: Tim Wrigley

ABSTRACT

Weight transfer in the golf swing is considered important in the coaching literature. However, scientific studies on weight transfer have been conflicting due to a number of limitations. This thesis examined weight transfer in the golf swing using more trials per golfer, Club Velocity at ball contact rather than handicap to indicate performance and more swing events at which weight position was quantified. Also, cluster analysis was used to identify if different swing styles exist. In study 1, 62 golfers performed ten simulated drives, hitting a golf ball into a net, while standing on two force plates. CP position relative to the feet (CPy%) was quantified at eight swing events identified from 200 Hz video. Cluster analysis identified two CPy% styles: 'Front Foot' style and 'Reverse' style. Both styles began with CPy% positioned evenly between the feet, moved to the back foot during backswing and then forward during early downswing. Beyond early downswing, the Front Foot group continued to move CPy% towards the front foot to ball contact, while the Reverse group moved CPy% towards the back foot to ball contact. Both styles occurred across skill levels from professional to high handicap golfers, indicating that neither style was a technical error. In study 2, group-based relationships between CP parameters and Club Velocity for each swing style was examined. For the Front Foot group, a larger CP range and a more rapid CP movement in downswing were associated with a larger Club Velocity at ball contact. For the Reverse group, positioning CP further from the back foot at late backswing and a more rapid CP transfer towards the back foot at ball contact was associated with a larger Club Velocity at ball contact. In study 3, individual-based analysis was conducted on five golfers performing 50 swings under the same test conditions. All golfers returned significant relationships between CP parameters and Club Velocity but these were individual specific. The most consistently related parameter was CP range, which was significantly related to Club Velocity for all golfers. Nonlinear techniques also were explored, with Poincare plots returning useful results for some golfers. In conclusion, analysis of weight transfer in the golf swing requires styles to be identified prior to any performance analysis. Individual-based analysis as well as group-based analysis is required to extract the most relevant information. Further, the use of more trials per golfer and more swing events should be employed in future studies.

List of Contents

Page

List of Tables

List of Figures

Page

This thesis is supported by the following publications

Study 1

Ball, K. and Best, R.J. (in press). Different weight transfer patterns in golf I: cluster analysis. *Journal of Sports Sciences*.

Study 2

Ball, K. and Best, R.J. (in press). Different weight transfer patterns in golf II: inter-individual analysis. *Journal of Sports Sciences*.

Acknowledgements

Russell Best for the guidance and imparting the ability for deeper thinking
Tim Wrigley for all things technical
Steve Dowlan and Dean Brown for the great sports science committee
Simon Taylor for the assistance with non-linear analysis
John Baker for the proof reading and counsel
Rob and Catharine for the support in so many ways
A very special thanks to mum and dad for giving me the opportunities and means over all these years to follow the path I have chosen – I will follow in your footsteps
To my sons Sam, Ben and Jack – thanks for making it harder and so much more enjoyable.

And the greatest thanks to my partner, Anne. This was a joint effort although you did the harder part. Without your incredible support this thesis would not have been completed.

I, **KEVIN BALL**, declare that the PhD thesis entitled **WEIGHT TRANSFER STYLES IN THE GOLF SWING: INDIVIDUAL AND GROUP ANALYSIS** is no more than 100,000 words in length, exclusive of tables, figures, appendices, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

Signature Date

CHAPTER 1: INTRODUCTION

Weight transfer in golf is a coaching term that refers to the movement of bodyweight and the resulting effect on the distribution of forces between the feet during the golf swing. Leadbetter (1995) describes this movement as weight balanced between the feet at address, moving towards the back foot during backswing and weight moving to the front foot in downswing and follow through. Golf coaching literature has stressed the importance of correct weight transfer in the golf swing, suggesting it is vitally important in developing momentum in the golf swing (Leadbetter, 1993) and is crucial to all golf shots in terms of length of hit, particularly for the driver (Norman, 1995).

In spite of the coaching emphasis on weight transfer, support for its importance in the scientific literature has been weak and often conflicting. In examining golfers of different skill levels, Wallace *et al*. (1990, N = 2) and Koenig *et al*. (1993, N = 14) found low handicap golfers produce greater amplitude and speed of weight transfer during the swing as compared with high handicap golfers. Conversely, Richards *et al.* (1985, N = 20) showed no differences in weight transfer between high and low handicap golfers. In examining weight transfer and performance, Robinson (1994, N = 30) found significant associations between weight transfer and club velocity at ball contact using golfers from a wide range of skill levels, while Mason *et al.* (1995, N = 64) found no association between weight transfer and club velocity at ball contact in

golfers with single figure handicaps. The importance, then, of weight transfer and forces at the feet in the golf swing remains unresolved.
There are a number of factors that may have influenced or limited previous weight transfer studies.

First, as noted by Wallace *et al.* (1994), the use of handicap-based groups to examine swing differences is not wholly appropriate since handicap does not always relate to swing quality. Rather, a measure of the quality or performance of the individual swing being tested is a more appropriate method to evaluate performance factors in the swing.

Second, the use of only one to four trials, the number of trials used for most golf weight transfer studies, may not be enough to indicate a typical/average swing. Bates *et al.* (1983), for example, have shown it can take up to eight trials to establish stable means in ground reaction forces in other sports.

Third, the use of only a small number of events in the swing (e.g. top of backswing, ball contact) at which the weight transfer data has been quantified reduces studies to more manageable levels, but loses all data/information between the chosen events. More events will provide more information, which may prove to be important in the golf swing.

Fourth, even though there is repeated comment on the different swing styles exhibited by different professionals, no studies have attempted to objectively identify and account for possible weight transfer styles. The idea of a single ideal model for a particular movement has been questioned in a number of activities (e.g. in shooting by Zatsiorsky and Aktov, 1990; in running by Dufek *et al*., 1995) and different swing strategies or styles might be expected in the golf swing as well. If they exist, different swing styles need to be examined separately, as important performance characteristics may differ between the styles. Further, if different styles exist, grouping all golfers together without accounting for the different styles would affect statistical analyses, with the influence of groups within the data generating type 1 or type 2 errors.

Finally, there has been no individual-based analysis of weight transfer in the golf swing. In other sports, significant results have been found on an individual basis but when the same individuals were used in group-based analysis, no effects were evident (e.g. in shooting: Ball *et al.*, 2003a, 2003b). This information would have been lost if only group-based analysis had been performed.

The aim of this study was to examine the relationship between performance and weight transfer in the golf swing on a group and an individual basis. This examination used club velocity at ball contact to indicate performance, used ten trials to establish each player's mean, and used eight swing events. Further, cluster analysis was used to search for different swing styles prior to performance analysis.

CHAPTER 2: LITERATURE REVIEW

2.1 WEIGHT TRANSFER IN THE GOLF SWING

This literature review contains discussion of weight transfer between the feet in golf. This has been limited to studies using the driver with the exception of Carlsoo (1967) who only tested the 5-iron. Weight transfer has been used to describe the different measures used in the literature including centre of pressure (CP), vertical force (Fz) and centre of vertical forces (COV). Briefly, Fz% has been measured by placing the back foot on one force plate and the front foot on another plate. The relative weight on each foot is the percentage of Fz on the front foot divided by the total Fz. CP is the point of intersection between the force vector and a defined horizontal plane, usually the ground surface. It is very similar to Fz% (Fz% and CPy% correlated strongly in this study: $r = 0.99$, $p < 0.001$, $N = 62$). COV is Fz% using one force plate. As both feet were on the same force plate, the feet needed to be digitised to determine the location of Fz relative to the feet. These measures are discussed in Appendix A.

2.1.1 What is weight transfer?

Weight transfer in the golf swing is a coaching term used to describe movement of bodyweight between the feet during the swing. Leadbetter (1995) describes the sequence of weight transfer (represented graphically in figure 2.1 with swing events presented in figure 2.2) as:

- Weight evenly balanced between the feet at address
- During the backswing, weight is initially moved towards the back foot
- Just before the top of backswing, backwards movement ceases and weight begins to move towards the front foot
- During the early downswing phase, weight is rapidly transferred towards the front foot
- At ball contact and follow through, weight is positioned on the front foot

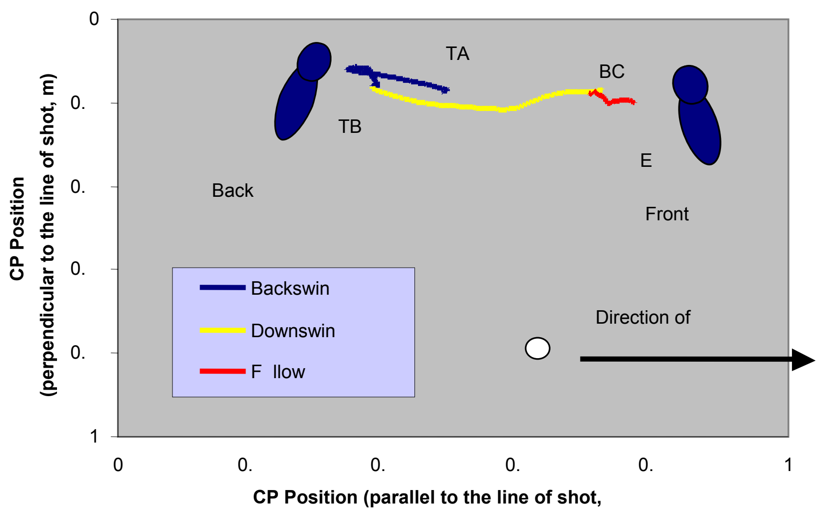


Figure 2.1: Example weight transfer profile, as indicated by centre of pressure (CP – line between the feet), during the golf swing (TA, TB, BC and EF are described in figure 2.2).

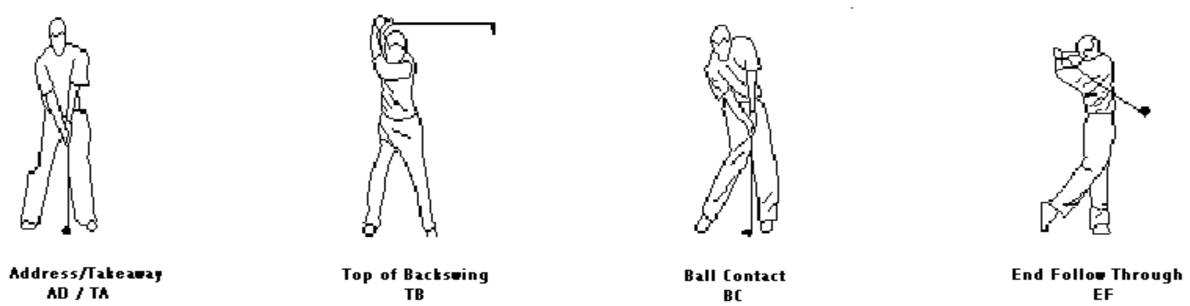


Figure 2.2: Swing events used in figure 2.1.

However, while weight transfer in the golf swing is frequently discussed in the coaching literature, the definition of weight transfer is somewhat unclear. Most coaching literature describes weight as the relative amount of body weight on each foot. For example, Newell and Foston (1995) suggested that weight should be felt on the front foot at ball contact, while Norman (1995) recommended that the position of weight should be evenly balanced on each foot at address and on the front foot at ball contact. Similarly, other authors refer to positioning weight on the front or back foot during the swing (e.g. Leadbetter, 1993; Leadbetter, 1995; Grant *et al.*, 1996).

Conversely, in biomechanics, 'weight' is defined as mass (m) multiplied by acceleration due to gravity (g), or Weight = m*g. Weight is considered to act through the centre of mass of the body (CM; e.g. Kreighbaum and Barthels, 1985). Weight transfer in biomechanical terms, then, might refer to the movement of the CM during the swing or the position of CM (transposed onto the horizontal plane) relative to the feet.

The relationship between the coaching weight transfer (weight under each foot) and the biomechanical definition of weight transfer (CM) has been examined in the scientific literature. Cooper *et al.* (1974) reported that CM and Fz% (relative vertical force, or weight, under each foot) followed a similar path and produced similar values during the golf swing. Five elite golfers (college scholarship holders) performed swings with the driver, 3-iron and 7-iron. CM was calculated from digitized data (21 body landmarks) from address to the end of follow through. It was then projected to the horizontal XY plane at the level of the feet and normalized to foot position. As well, two force plates, one under each foot, measured the force under the front and back foot from which Fz% was calculated (percent of Fz on the front foot during the swing; i.e. 100% = front foot). Cooper *et al.* reported that there was a 'great similarity' between CM and Fz% from AD to BC for all clubs, at which point both indicated weight was approximately 75% towards the front foot. At mid follow through, Fz% tended to be moved towards the back foot (50%) before again moving to the Front Foot (70%), while CM continued to move towards the front foot at both events (no values reported).

Conversely, Mason *et al.* (1990) reported that the paths of the CM and centre of pressure (CP, CP% when expressed relative to the distance between the feet) differed at the top of backswing and in downswing. Mason *et al.* calculated CM from digitized data (similar to Cooper *et al.*, 1974) from 12 golfers (handicap < 10) performing a swing with the driver. CP was calculated from two force plates, one under each foot, and normalized to the digitized foot position to obtain CP% between the feet. In pilot work for this study, CP%, as used by Mason *et al.* (1990) and Fz%, as used by Cooper *et al.* (1974), were found to be very similar ($r = 0.99$, $p < 0.001$; see Appendix B) and so it is unlikely that the differences between studies can be explained by the different measures used.

CM and CP have been compared in other tasks, such as postural sway. Winter (1995) reported that CP in the anterior-posterior direction of body sway was in phase with CM movement, but with larger amplitude. In the medio-lateral direction, CM and CP were anti-phase. This highlights the potential differences that can exist between the two measures. A number of attempts to approximate CM and CP or other force plate measures have been made. For example, Lafond *et al.* (2004) reported the zero to zero double integration method applied to CP was comparable to the kinematic (direct) method of determining CM sway. However, for this and other methods used to approximate CM require task specific considerations to be effective and cannot necessarily be used for other tasks.

While differences might exist between the coaching and biomechanical definitions of weight transfer, there is an important distinction to note. Weight transfer as used in the golf coaching literature is a coaching term and is used to describe a particular aspect of the skill. As such, the basis of the term in the context of golf is in coaching and not in biomechanics. The measurement of weight transfer, then, should reflect this understanding of weight transfer rather than a scientific definition. As such, the use of CP% or Fz% is more appropriate than the measurement of the CM to explore weight transfer as it relates to the relative size of the forces beneath each foot.

2.1.2 Measurement of weight transfer

Table 2.1 reports the measures used in studies examining weight transfer. In all cases, the measures used have reflected the coaching emphasis of weight under each foot rather than CM motion with all using either CP% or Fz% (or slight variations of these measures) to define weight transfer. The term ‘parallel to the line of shot’ is used in reference to weight transfer between the feet. Weight transfer from heel to toe of each foot is referred to by ‘perpendicular to the line of shot’.

Table 2.1: Type of measurement used to indicate weight transfer parallel to the line of shot in golf studies (arranged by measure used).

Researcher	Measure Used	Notes
Cooper *et al.* (1974)	Fz%	Two force plates (one under each foot)
Williams and Cavanagh (1983)	Fz%	Two force plates (one under each foot)
Koenig *et al.* (1993)	Fz%	Two force plates (one under each foot)
Barrentine *et al.* (1994)	Fz%	Fz% used to describe between feet movement
Robinson (1994)	Fz (but reported in kg)	Two force plates (one under each foot) - Fz reported in kg and normalised to body mass
Carlsoo (1967)	Fz (but reported in kg)	Two force plates (one under each foot) - Fz reported in kg for right and left feet
Richards *et al.* (1985)	COV (Fz%)	One force plate - Fz% relative to digitised foot position
Wallace *et al.* (1990)	CP%	Two pressure plates (one under each foot)
Mason *et al.* (1990)	CP%	Two force plates (one under each foot) – CP relative to digitised foot position
Mason *et al.* (1995)	CP%	Two force plates (one under each foot) – CP relative to digitised foot position
Neal (1998)	CP%	One force plate (one under each foot) – CP relative to digitised foot position
Koslow (1994)	'weight shift patterns'	Sporttech swing analysis system – no further details provided in the study

Note: Williams and Cavanagh (1983), Koenig *et al.* (1993) and Barrentine *et al.* (1994) used CP as well as Fz% but CP was used for heel to toe analysis only.
COV = centre of vertical forces = Fz%.

Seven studies used vertical forces to indicate weight transfer in the golf swing. In four studies (the first four in table 2.1), Fz was evaluated using two force plates, one under each foot. Fz% was then calculated from this data by dividing the Fz under the front foot by the total Fz (equation 2.1). In one study, two force plates were used and the weight transfer measure was defined as vertical force but values were reported in kilograms under each foot (Carlsoo, 1967). While this is not a unit of force, the results differ only by a constant (acceleration due to gravity). Robinson (1994) also used two force plates and also reported vertical force in kilograms, as well as normalised to body mass under each foot. It was not clear to this researcher if this was performed using body mass as measured during quiet stance (standing still) or, as used in other studies, by total vertical force divided by g at any instant. Using the values reported by Carlsoo (1967), vertical force varied between 0.9 to 1.3 times body mass in different parts of the swing so this distinction is important. Using body mass measured during quiet stance will provide different values of normalised Fz than if the total instantaneous vertical force was used.

$$\text{Position of Fz relative to the feet} = \frac{\text{Fz (under front foot)}}{\text{Fz (under front foot)} + \text{Fz (under back foot)}}$$

Equation 2.1

One study (Richards *et al*., 1985) used a measure termed centre of vertical forces (COV), defined as the point at which the vertical force vector (Fz) intersected the horizontal plane at ground level. The toe and heel of each foot was digitised from video footage captured while the golfer was in the address position. The midpoint of each foot was calculated and used to indicate 0% (back foot) and 100% (front foot) of the distance between the back and front foot. The position of the COV was then calculated relative to the feet. This process is represented graphically in figure 2.3.

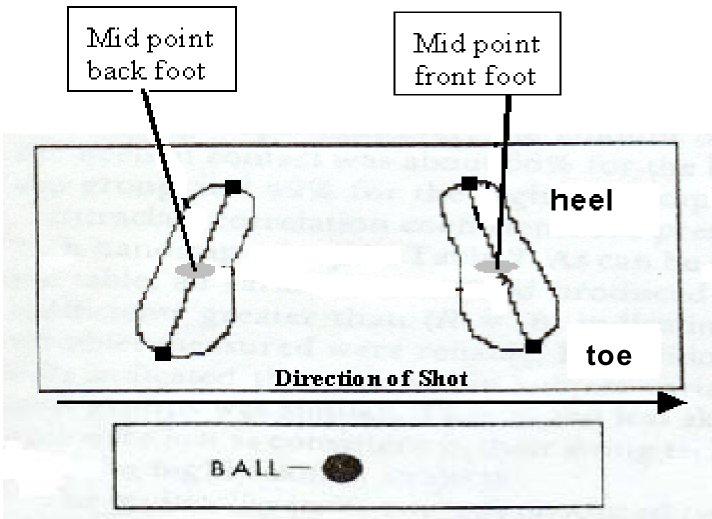


Figure 2.3: Determination of mid-foot position from Richards *et al.* (1985)

Four studies used CP to indicate weight transfer. In three of these studies, CP was calculated from force data measured using one or two force plates. Foot position was obtained using similar methods to those used by Richards *et al.* (1985) and CP location was expressed as a percentage of the distance between the feet. The fourth study using CP (Wallace *et al.*, 1990) did not use foot digitising. Wallace *et al.* reported using a 'Musgrave' foot plate under each foot and used a similar rationale to Fz% calculation was used to obtain CP% (pressure on front foot plate divided by the sum of pressure on the front and back foot plate, equation 2.2).

$$\text{CP (between the feet)} = \frac{\text{Pressure (under front foot)}}{\text{Pressure (under front foot)} + \text{Pressure (under back foot)}}$$

Equation 2.2

The final study reported in table 2.1 did not report the type of measure used. Koslow (1994) used a Sporttech Swing Analyser, which included a force plate that measured "weight shift patterns" during the swing. Koslow reported the results as a percentage of body weight placed on the back foot at different swing events but did not detail how this was performed. This researcher was unable to find details of this Sporttech system either in the literature or in golf equipment and internet searches.

2.1.3 Description of weight transfer in the scientific literature

The common coaching model described in section 2.1.1 has been supported in the scientific literature (e.g. Carlsoo, 1967; Cooper *et al.*, 1974; Williams and Cavanagh, 1983). However, this support has largely taken the form of qualitative assessment of quantified data (Fz% or CP). For example, Williams and Cavanagh (1983) reporting that weight (Fz%) was positioned approximately in the middle of the feet at address, moved towards the back foot during backswing, then rapidly forward in the downswing towards the front foot. A similar description was provided by Cooper *et al.* (1974) and Koenig *et al.* (1993), also using Fz%.

Empirical evidence for the weight transfer pattern during the swing is difficult to obtain as the reporting of data in the scientific literature has been inconsistent and incomplete. Only seven of the 12 weight transfer studies in the literature reported data relating to the position of weight at different swing events. Of these seven, only three studies reported enough data to evaluate an overall pattern of weight transfer from address to ball contact or the end of follow through (Carlsoo, 1967; Wallace *et al.*, 1990; Koslow, 1994). Table 2.2 details all reported data for weight transfer between the feet in the literature. Figure 2.4 shows the swing events used in different studies at which data has been reported (i.e. AD, TA etc. in table 2.2: these abbreviations are used throughout this thesis). Figure 2.5 shows another event defined by Robinson (1994) as when the forearm is in the horizontal plane in downswing. No data was reported for this event.

Table 2.2: Weight transfer data reported in the literature (between the feet - swing events are defined in figures 2.4 and 2.5)

Study	Group tested	Measure	AD	TA	MB*	TB	MD*	BC	MF*	EF	Min	Max
			Swing Events									
Cooper *et al.* (1974)	Elite College players (*N*=5)	Fz%						75	50	70		
Richards *et al.* (1985)	Low Handicap (<10, *N*=10)	COV				28		96			17	105
	High Handicap (>20, *N*=10)					22		81			15	98
Wallace *et al.* (1990)	Low Handicap (*N*=1)	CP%	63		53	27	68	82		90		
	High Handicap (*N*=1)		49		42	31	47	67		77		
Koenig *et al.* (1993)	Handicap (0-20, *N*=14)	Fz%		55		35					20	
Robinson (1994)	Professional (*N*=10)	Fz%		49								
	Amateur (Handicap 0-20, *N*=20)			58								
Koslow (1994) **	Novice Correct weight shift (*N*=5)	"weight shift patterns"	48			27		62				
	Novice Abbreviated weight shift (*N*=17)		49			39		43				
	Novice Reverse weight shift (*N*=8)		51			60		36				
Carlsoo (1967)	See table 2.3											
Williams and Cavanagh (1983)	No values reported											
Mason *et al.* (1990)	No values reported											
Barrentine *et al.* (1994)	No values reported											
Mason *et al.* (1995)	No values reported											
Neal (1998)	No values reported											

All values expressed as a percentage relative to the feet (0% = back foot, 100% = front foot). Transformed from the data presented in each study if required to allow for direct comparison between studies.

* MB, MD and MF have not been well defined in these studies

** Koslow used set-up (assumed to be AD), top of swing (assumed to be TB) and BC and described the measure used only as 'weight shift patterns' with no further explanation.

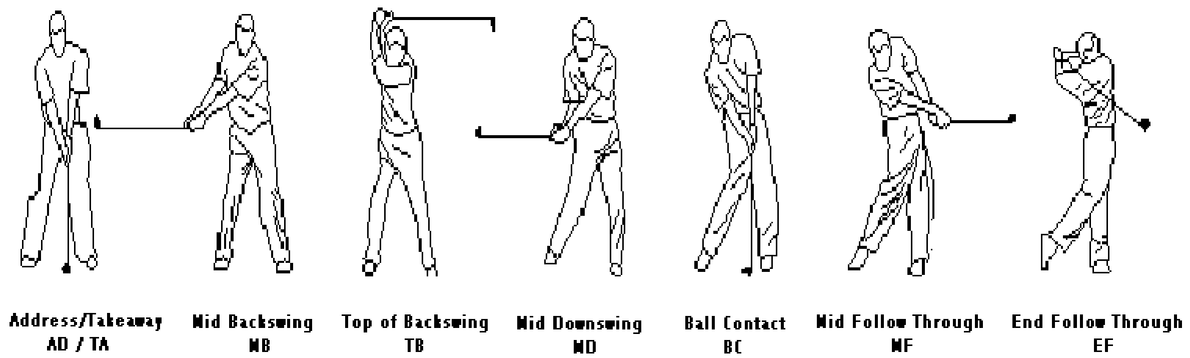


Figure 2.4: Swing events used in golf studies (abbreviations used in this thesis).

Note: MB, MD and MF have not been well defined in the literature. The diagrams here are based on the definition for this study and have been included to illustrate their approximate position in the swing

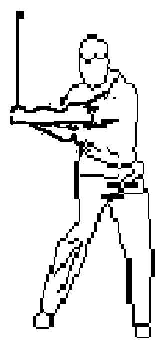


Figure 2.5: The swing event defined as "forearm horizontal in downswing" (FH) as used by Robinson (1994).

In the studies that have reported data for the entire swing, the coaching model of weight transfer as defined in section 2.1.1 has been supported. Figure 2.6 presents the weight transfer patterns from Wallace *et al.* (1990) with weight at AD being located balanced between the feet for the low handicap golfer, moving towards the back foot in backswing for both golfers (MB and TB) and to the front foot in downswing and follow through for both golfers (MD, BC, EF). Koslow (1994) also reported this pattern occurring, but only for five of the N = 30 novice golfers tested.

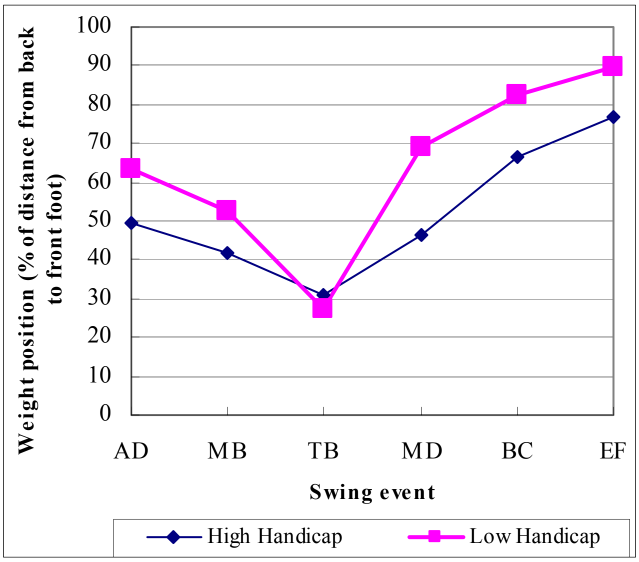


Figure 2.6: Comparison of weight transfer for a high and low handicap player during the golf swing (from Wallace *et al.*, 1990).

Carlsoo (1967) measured and reported the position of weight at eight swing events. While the swing event definitions were not clear, there is enough data to indicate the coaching model of weight transfer. Table 2.3 reports the events defined by Carlsoo (described in the article as phases rather than events) and Fz data at these events. Event descriptions are taken directly from the text of the Carlsoo article. However, it should be noted, while Carlsoo, as well as Wallace et al. (1990) and Koslow (1994) have reported the coaching-based weight transfer pattern, in total this represents only

N = 8 golfers (Carlsoo, 1967: 1 golfer, Wallace *et al.*, 1990: 2 golfers, Koslow, 1994: 5 golfers). Clearly there is a need to measure and report weight transfer data in the scientific literature.

Table 2.3: Swing event definitions and Fz% calculated from data presented by Carlsoo (1967).

	Fz (kg) on front foot	Fz (kg) on back foot	% to front foot
At the beginning of the address	42	42	50
At the end of address	48.5	35.5	58
At the start of backswing	34	46.5	42
At the beginning of the backswing	29	53	35
At the top of the backswing	49.5	34.5	59
Immediately before impact	88	24.5	78
The impact	74	24	76
Immediately after impact	54	21	72

Note: Carlsoo reported Fz data in kg rather than Newtons but this does not change the Fz% calculation.

While empirical data has been inconsistently reported, the general weight transfer pattern of balanced at address, on the back foot at the top of backswing and on the front foot at ball contact is supported if data from all studies is examined together. Figure 2.7 shows values (Fz%, CP%, COV) reported in the literature for AD, TB and BC (Koslow, 1994, excluded for this comparison as only novice golfers were used). The position of weight at AD has been reported in two studies and ranged from 49% to 63% (Carlsoo, 1967; Wallace *et al.*, 1990). This supports the ‘approximately equal’ term used in coaching literature. The 63% reported by Wallace *et al.* (1990) might be considered somewhat divergent from ‘balanced’, although it was produced by a high handicap golfer. At TB, reported values have ranged from 22% to 35% in three studies (Richards *et al.*, 1985; Wallace *et al.*, 1990; Koenig *et al.*, 1993) suggesting that weight is positioned towards the back foot at TB. As well, the smaller values in TB compared with AD indicated that weight moved towards the back foot in backswing. At BC, reported values range from 67% to 95% in four studies (Carlsoo, 1967; Cooper *et al.*, 1974; Richards *et al.*, 1985; Wallace *et al.*, 1990), suggesting that weight is positioned predominantly on the front foot at BC, although the 67% reported by Wallace *et al.* (1990) for a high handicap golfer might be considered low. As well, the larger values compared to those reported at TB indicate that weight had moved towards the front foot during downswing.

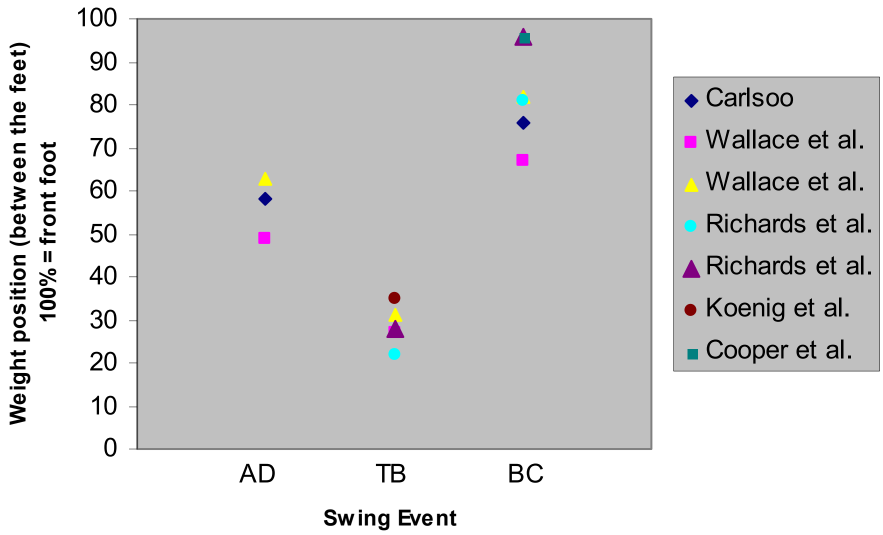


Figure 2.7: Weight transfer values reported in the literature for AD, TB and BC.

While support is provided at these swing events, a limitation of these studies is the small number of swing events at which weight position has been reported. It is not known where weight is positioned between these key events and so useful data might have been missed. This was highlighted in Cooper *et al.* (1974) who reported Fz% at BC was 75%, moving to 50% at mid follow through and then to 70% by the end of follow through. This indicated a shift towards the back foot between ball contact and MF, a technical trait not discussed in coaching texts. No comment is made of this data in the Cooper *et al.* article, nor has any other study reported data at both MF and EF for a full comparison. However, Carlsoo (1967) reported the Fz% was lower ‘immediately after impact’ (Fz% = 72%) compared with BC (Fz% = 76%) for an elite golfer. Although the difference is only small, it supports the possibility that weight moves towards the back foot after BC and is moving towards the back foot at BC for some golfers. While weight was positioned further towards the front foot at EF compared with BC for Wallace *et al.* (1990) no measure was taken between BC and EF (e.g. immediately after BC as used by Carlsoo, 1967) and so it cannot be evaluated if this pattern occurred for these golfers. While the Cooper *et al.* (1974) finding was interesting on its own, the fact that the players tested were elite would strongly support more thorough investigation of the weight transfer pattern between TB and BC and between BC and EF.

2.1.4 Importance of weight transfer in the golf swing

Golf coaching literature has stressed the importance of correct weight transfer in the golf swing (e.g. Grant *et al.*, 1996; Leadbetter, 1995; Frank, 1994). Leadbetter (1993) noted that weight transfer is vitally important in developing momentum in the golf swing and is a prerequisite for a solid, powerful swing. Leadbetter also highlighted two weight transfer flaws which affect golfers: the reverse pivot, where the weight moves to the front foot in backswing and to the back foot in forward swing; and the lack of weight transfer, where the weight remains on the back foot. Norman (1995) suggested that weight transfer is crucial to all golf shots in terms of length of hit particularly for the driver. Madonna (2001) included weight transfer as one of four 'absolutes' in the golf swing. These absolutes were defined as skill factors that are essential to the production of a good swing. Specifically, Madonna (2001) suggested that weight should move to the target leg (the leg closest to the target, referred to in this study as the front foot) after TB while the arms come down to the trailing side of the body.

However, not all the coaching literature has supported the importance of weight transfer in the golf swing. Cooke (1987) suggested that the body should move about a vertical axis and, as such, no weight transfer should exist. Rae *et al.* (2001) suggested that weight transfer is not important in itself, but is necessary to allow the body to rotate optimally. Clearly, even within the coaching literature, there is some contention as to its importance.

2.1.4.1 Position of weight at different swing events

Most coaching literature has stressed where weight should be positioned at different events in the swing (e.g. Leadbetter, 1993). However, only two of the five scientific studies that have reported examining these factors found either differences in weight position between skill levels or that weight position was related to a performance measure. Further, no two similar studies have reported significant results for position of weight at the same swing event. For example, while Robinson (1994) found a relationship between weight position at TA and club velocity at ball contact, Mason *et al.* (1995) found no significant effect for a similar analysis at $p < 0.05$. In comparing handicap groups, Wallace *et al.* (1990) found one high and one low handicap golfer differed at TA, MB, MD, BC and EF. However, Richards *et al.* (1985) found no statistical difference between a high and a low handicap group ($N = 10$, $p < 0.05$) at TB or BC, although medium effect sizes existed (calculated by this researcher and discussed below).

The only event that has been indicated as important in examining weight transfer and performance (TA, Robinson, 1994) has not been supported in other studies, providing conflicting and as a result inconclusive data. Robinson, using 30 golfers from a wide range of skills (professional to high handicap), found an association between the vertical force on the back foot at takeaway and club velocity at ball contact ($r = 0.45$, $p = 0.02$). This indicated that a larger Fz on the back foot, or positioning weight further towards the back foot, was associated with a larger club velocity. Robinson also reported that professional players tended to adopt a more balanced position of 51% on the back foot, while amateur players adopted a stance with weight further towards the front foot (only 42% on the back foot). No statistical analysis was reported for this comparison. These values were reported relative to the front foot in table 2.2 to make them consistent with other studies but have been reported here as they were presented by Robinson. Mason *et al.* (1995) found no association between club velocity at ball contact and the position of CP between the feet at TA for 64 apprentice professional golfers. It is possible that the small range of skill levels in the golfers tested influenced this result as a more homogenous sample may lead to a smaller effect size (e.g. Coleman, 1999). No other study has reported the relationship between the position of weight at TA and performance.

The comparison of weight position at TA between skill levels has been limited and might also be conflicting. Wallace *et al.* (1990) reported finding significant differences at TA between skill levels using one low and one high handicap golfer. Wallace *et al.* reported that the low handicap golfer positioned CP closer to the front foot at TA (low = 63%; high = 49%; one way ANOVA, $N = 10$ shots, $p < 0.001$: no other statistics such as effect size or F-ratios were reported for a more precise evaluation). It should be noted that both golfers used in the Wallace *et al.* study would have been part of the amateur group in the Robinson (1994) study, so for this reason alone the group comparison aspect of these two studies cannot be compared. No other study has reported differences between skill levels for TA.

Only one study has reported finding statistical support for the importance of weight position at swing events other than at TA. Wallace *et al.* (1990) examined one high handicap and one low handicap golfer performing 10 swings with the driver. CP between the feet was quantified at six swing events: TA, MB, TB, MD, BC and EF. CP at all swing events except TB was significantly different between the golfers. However, these findings have not been supported by other research. Robinson (1994) found no significant relationships between club velocity and Fz% at TB, FH, BC and MF (MF was defined as 0.25 s after BC). Robinson expressed surprise in finding no association between club velocity and the position of weight at ball contact, noting that moving the weight to the front foot at ball contact is a swing characteristic emphasised in golf instruction and swing measurement devices. Similarly, Mason *et al.* (1995) found no significant associations between club velocity and CP between the feet at TA, TB, BC and EF. As well, Richards *et al.* (1985) found that the position of centre of vertical forces (COV) at TB and BC did not differ between high (> 20) and low (< 10) handicap players.

It is important to note that low statistical power and limited generalisability due to small subject numbers existed in all weight transfer studies. Wallace *et al.* (1990) used only $N = 2$ subjects making it inappropriate to generalise the study's findings. Also affected by low N, the Robinson (1994; $N = 30$) study possessed low statistical power. To achieve the 80% power recommended by Cohen (1968; i.e. 80% chance that a significant relationship will be found if it exists), Robinson (1994) would have needed to find large effect sizes among the relationships tested ($r > 0.5$; for the study's parameters of $p = 0.05$ and $N = 30$). As the study was presented, the minimum detectable effect was $r = 0.35$, and so small effects were not detectable (using levels defined by Cohen, 1968; small effect: $r > 0.2$, moderate effect: $r > 0.3$, large effect: $r > 0.5$).

Richards *et al.* (1985) offset the loss in statistical power due to low N by increasing the expected effect size. Richards *et al.* compared groups that were further apart in skill level than all other studies (Low Handicap 0-10; High Handicap 20+). However, the effect size was decreased due to a large variance in the high handicap group and so the attempt to increase power by choosing skill levels that differed more than in other studies was offset by the less skilled group being highly variable in performance. For the Richards *et al.* study, effect sizes were medium (calculated by this researcher using Cohen 1988, small: $d = 0.2$, medium: $d = 0.5$, large: $d = 0.8$), for the difference between groups at TB (effect size $d = 0.51$) and BC (effect size $d = 0.73$). With a medium effect size and alpha set at $p = 0.05$, more than 25 subjects in each group were needed to achieve 80% power. With only $N = 10$ as used in Richards *et al.*, even a large effect size (effect size $d > 0.8$) would have possessed only 40% power (i.e. there is less than an even chance of finding a significant result if it exists).

Mason *et al.* (1995) collected a larger sample *(N* = 64) but the narrow range of skill levels reduced statistical power (PGA apprentice professionals only). Small effect sizes can be expected in elite level groups as the difference between subjects will be small (e.g. Ball *et al.*, 2003a; Ball *et al.*, 2003b; Coleman, 1999), and this will in turn reduce statistical power (e.g. Aron and Aron, 1999). The power to detect small effects in the Mason *et al.* (1995) study was only 36%. Further, even with $N = 64$, some small effects would not have been detected in the study as, at $p < 0.05$, an r-value of 0.24 was required for significance.

2.1.4.2 Range of weight transfer

One technical factor that has had support from more than one study is the range of weight transfer during the swing. Wallace *et al.* (1990, $N = 2$) reported that a low handicap player (handicap = 6) showed greater amplitude of weight transfer than a high handicap player (handicap = 24). As can be noted in figure 2.8 (repeated from section 2.1.3), the low handicap golfer moves weight further towards the back foot at TB (although not significant between golfers) and further towards the front foot at BC and EF (both significant between golfers, $p < 0.05$). Koenig *et al.* (1993) also reported that low handicap players exhibited greater amplitude of weight transfer than high handicappers in comparing three skill levels *(N* = 14, handicap 0-7, 8-14, 15+). However, dissimilar to Wallace et al. (1990) who showed a difference between skill levels at BC and EF, Koenig *et al.* (1993) reported the differences existed at TB, with higher handicap golfers tending to show a more balanced position at TB (i.e. weight was evenly balanced between the feet) than the low handicap golfers who positioned weight near the back foot. As such, differences between studies have once again been highlighted.

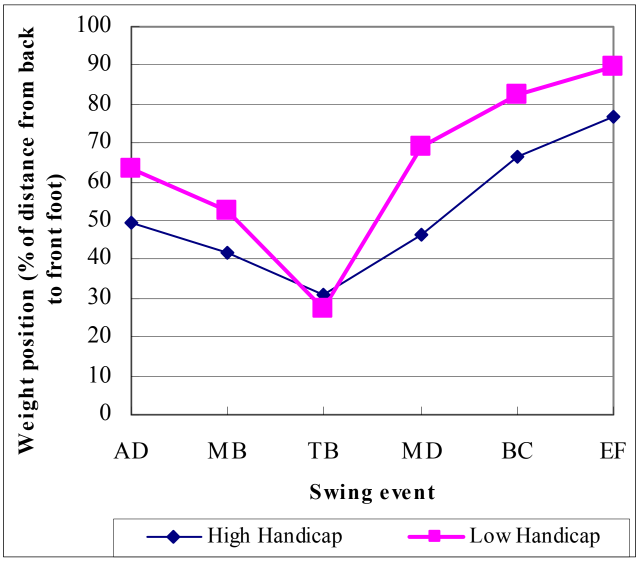


Figure 2.8: Comparison of weight transfer for a high and low handicap player during the golf swing (from Wallace *et al.*, 1990).

While Richards *et al.* (1985) did not analyse range specifically, enough data was presented for this parameter to be evaluated. Richards *et al.* reported weight position at TB and BC, as well as maximum (nearest the front foot) and minimum (nearest the back foot) weight positions for a low handicap group (handicap < 10, N = 10) and high handicap group (handicap > 20, N = 10). Table 2.4 reports the ranges calculated from this data. For all range calculations, low handicap golfers exhibited larger ranges compared to high handicap golfers. It is not possible to perform comparative statistics on this data, as standard deviations cannot be calculated for these parameters.

Table 2.4: COV% ranges calculated from data presented by Richards *et al.* (1985).

Group	TB-BC	Min-BC	Min-Max
Low Handicap (*N* = 10)	68	79	89
High Handicap (*N* = 10)	59	66	84
Difference	9	13	5

Min = most backward point of COV (i.e. nearest the back foot)
Max = most forward point of COV (i.e. nearest the front foot)

While both Wallace *et al.* (1990) and Koenig *et al.* (1993) reported differences between skill levels in golfers for weight transfer range, no statistical support for this relationship was presented. The difference in range of weight transfer is evident in figure 2.8 from Wallace *et al.* (1990) but no specific analysis was performed on range data. Koenig *et al.* did not present empirical evidence to support the comments that low handicap players exhibited greater range of weight transfer compared with high handicap golfers. As discussed in the previous paragraph, Richards *et al.* (1985) did not analyse range of weight transfer. No other study has reported examining weight transfer range. This lack of statistical support for range of weight transfer will be addressed in this study.

2.1.4.3 *Velocity of weight transfer*

Another finding that has had support relates to the rate or velocity of weight transfer (Wallace *et al.*, 1990; Koenig *et al.*, 1993; Robinson, 1994). Robinson (1994) found significant correlations between club velocity and the rate of change of Fz on the left (front) foot in kg/s and with the vertical force normalised to body mass (%/s) between TB and FH (i.e. Fz or Fz% at FH minus Fz or Fz% at TB divided by the time between events, table 2.5). For both relationships, more skilled golfers (professionals) produced greater rate of change of weight transfer than less skilled golfers (amateurs) although no statistical analysis of this data was reported.

Table 2.5: Parameter values and correlations between velocity of weight transfer and club velocity at ball contact from Robinson (1994)

Rate of change of Fz on the front foot between TB and FH	r	p	Professional ($N = 10$)	Amateur ($N = 20$)
kg/s	0.69	<0.001	158	119
Normalised (%/s)	0.61	<0.001	390	298

Unfortunately, Robinson (1994) is the only study to examine rate of weight transfer and report data for the findings. While similar observations were made by both Wallace *et al.* (1990) and Koenig *et al.* (1993), who reported that low handicap golfers exhibited a more rapid weight transfer early in downswing compared with high handicap golfers, neither study reported data to support these comments. As such, the support for velocity of weight transfer in scientific terms is poor and requires more thorough evaluation.

Also of interest in the Robinson (1994) study was that weight transfer velocity was important in the early downswing stages of the swing. This is a phase that no other study has examined. The limited number of swing events used in the literature will be discussed in section 2.1.5.

2.1.4.4 Summary of weight transfer findings

In summary, examination of weight transfer and the golf swing in the scientific literature has produced conflicting or inconclusive results. Despite the emphasis on the position of weight at different stages of the swing in the coaching literature, this has not been supported in the scientific literature. While weight transfer range and velocity have been observed as important, statistical support has been limited to one study only. More generally, the scientific literature examining weight transfer in the golf swing has suffered from low subject numbers, low statistical power, no individual-based analyses examining performance and a lack of detail in statistical analysis. There is a need for more thorough analysis of the weight transfer in the golf swing with increased statistical power and with strong statistical designs. These issues are addressed in this study.

2.1.5 Methodological issues in previous research

Other than the issues of low *N* and statistical power already discussed, a number of methodological issues exist in the literature. These issues may also have contributed to the conflicting and non-significant findings in examining weight transfer in the golf swing:

2.1.5.1 The use of handicap rather than swing performance measures

2.1.5.2 The low number of trials used to assess a typical swing

2.1.5.3 The low number of swing events used to assess weight position

2.1.5.4 The lack of examination of different swing styles between golfers

2.1.5.5 The lack of individual-based analyses

2.1.5.1 The use of handicap rather than swing performance measures

Four criterion measures have been used in golf weight transfer studies (table 2.6).

Table 2.6: Performance criterion measures used in golf weight transfer studies.

Researcher	Criterion measure
Williams and Cavanagh (1983) Richards *et al.* (1985) Wallace *et al.* (1990) Koenig *et al.* (1993) Barrentine *et al.* (1994)	Handicap group Handicap group Handicap group Handicap group Handicap group
Robinson (1994) Mason *et al.* (1995)	Club Velocity Club Velocity
Koslow (1994) Neal (1998)	Swing style Swing style
Carlsoo (1967) Cooper *et al.* (1974) Mason *et al.* (1990)	Elite level player – descriptive only Elite level players – descriptive only Comparison of CP and CG

Of the twelve studies reviewed, five used the golfer's handicap to form groups and to compare weight transfer characteristics. The rationale behind this measure is that lower handicap golfers are expected to possess better swings. Weight transfer characteristics that differ between low and high handicap golfers might indicate important aspects or desirable characteristics of weight transfer. However, as noted by Wallace *et al.* (1994), the use of handicap-based groups to examine swing differences is not wholly appropriate since handicap does not necessarily relate to swing quality. Handicap, rather, is a measure that indicates the golfer's skill in all aspects of golf, including driving, putting, chipping and course management. Further, in three of the five studies using handicap groups, cut-offs used seem arbitrary and by necessity may have separated golfers who differ by only 1 handicap point. Koenig *et al.* (1993) and Williams and Cavanagh (1983) divided groups into handicaps of 0-7, 8-14 and 14+, while Barrentine *et al.* (1994) used 0 - 15 and 15 + to denote low and high handicap golfers. As well, there is little consistency between these studies with low handicap groups, for example, being defined at 0-7 (Koenig *et al.*, 1993; Williams and Cavanagh, 1983), 0-15 (Barrentine *et al.*, 1994) and 0-10 (Richards *et al.*, 1985). This makes it difficult to compare studies as groups defined as low and high handicap differ between studies. While this criterion has merit in assessing what might be important to success in the golf swing by examining what better (lower handicap) golfers do, information can be missed because handicap does not relate directly to the golf swing. As well, it offers little information for the low handicap golfer or elite player who is effectively being used as the ‘ideal’ model. For example, a golfer with a handicap of five is not offered any information on how to improve their swing in any of these studies, as the assumption in this type of testing is that the low handicap group exhibits the desirable characteristics already while the high handicap groups do not.

The use of swing characteristics to indicate performance has been employed by only two studies. In both cases, club velocity at ball contact was used. Club velocity is related to ball speed, which in turn is related to distance, which is regarded as a good performance indicator for the golf swing. While distance and accuracy are the indicators of performance on the golf course, the laboratory environment and lack of available measurement devices for ball flight characteristics have limited research to club velocity. This can also be seen as de-limiting the study, as the measure eliminates factors such as off centre impacts, which will affect ball flight but which may not be due to weight transfer factors (e.g. due to slight shift of the hands; Leadbetter, 1993). The use of club velocity rather than distance was an environmental issue for Mason *et al.* (1995), with testing being conducted inside a laboratory. Robinson (1994) tested on a driving range but distances were not reported.

The remaining five studies did not include a performance component. Both Neal (1998) and Koslow (1994) compared different styles of swing. Carlsoo (1967) and Cooper *et al.* (1974) were both descriptive studies. Mason *et al.* (1990) compared CP and CG.

2.1.5.2 The low number of trials used to assess a typical swing

The number of trials used to indicate a typical swing has varied between studies from 1 to 300 (table 2.7), with most using less than five trials. This is possibly an important measurement issue in weight transfer research as the use of only a few trials decreases the chances of collecting typical data (e.g. Mullineaux *et al*., 2000). Bates *et al.* (1983) found that ground reaction force (GRF) parameter means in running did not stabilise until eight trials had been collected. Briefly, Bates *et al.* calculated the mean, standard deviation and one quarter of this standard deviation of each GRF parameter for a large number of running trials. Then, beginning with the first trial, successive trial GRF values were added and the 'progressive' mean was calculated. The parameter was defined as stable once the change in the mean from trial to trial fell below one-quarter of the standard deviation calculated across all trials. This method was used in this study to determine the number of trials needed in testing (see Methods section 4.2.2.1).

Table 2.7: Number of trials used per individual golfer for studies reviewed.

	Number of trials used	Notes
Carlsoo (1967)	approx 300	Over 7 testing sessions
Cooper *et al.* (1974)	1	
Williams and Cavanagh (1983)	4	
Richards *et al.* (1985)	4	
Wallace *et al.* (1990)	10	Only 2 golfers tested
Mason *et al.* (1990)	1	
Koenig *et al.* (1993)	7	
Koslow (1994)	10	Not stated, but assumed by this researcher that all 10 trials were used in the analysis.
Barrentine *et al.* (1994)	3	3 trials with golf shoes, 3 with running shoes (not stated if analysed separately)
Robinson (1994)	1	Described as typical
Mason *et al.* (1995)	1	Described as typical
Neal (1998)	1	

Using different techniques, the reported number of trials needed for a consistent result to be returned has varied widely. As mentioned, Bates *et al.* (1983) reported requiring eight trials for means to stabilise in running. In other tasks, ten trials were required in walking (Giakas and Baltzopolous, 1997) and between 1 and 78 trials in hurdling (Salo *et al.*, 1997). Although no statistical analysis was performed to assess the point at which the mean stabilised, data presented by Best *et al.* (2000) indicated that mean minimum toe clearance (MTC) during the swing phase of normal gait did not stabilise even after 1000 trials. While MTC always remained within 3 mm (1 – 3.3 mm) and Best *et al.* commented that the data was stable from a general point of view, this variation in result needs to be evaluated in terms of the research question and the accuracy required. For example, if differences between subjects were of the order of 3 mm, then unstable mean values for individuals may be a factor in the study. With different results indicated for different skills (and for different parameters relating to the same skill) the evaluation of what constitutes a stable mean clearly needs to be specific to the measurement parameter and requires evaluation prior to testing.

The effect of the number of trials used to assess the golf swing in previous studies needs to be considered. For studies using handicap groups as criterion measure, if the swings used to represent each individual are not typical, then group means for each handicap group will not be typical. As such, relationships may not be detected (type 2 error), or random relationships may arise (type 1 error) in different studies. A similar issue exists when examining performance-based factors such as club velocity. If the measured swing for a particular golfer was atypical, then this will affect ensuing analyses and generate either type 1 or 2 errors.

2.1.5.3 The low number of swing events used to assess weight position

Much of the research in weight transfer has examined data at different swing events (although as mentioned this data is often not reported). Table 2.8 reports the swing events used in different studies.

Table 2.8: Swing events at which weight position was quantified in previous research.

	No. of swing events	AD	TA	MB	TB	FH	MD	BC	MF	EF	Max	Min
Studies providing statistical analysis												
Richards *et al.* (1985)	2				•			•			•	•
Wallace *et al.* (1990)	6		•	•	•		•	•		•		
Robinson (1994)	5		•		•	•		•		•		
Koslow (1994)	3		•		•			•				
Mason *et al.* (1995)	4		•		•			•		•		
Descriptive studies												
Koenig *et al.* (1993)	2		•		•							•
Carlsoo (1967)*	8	•	•		•			•				
Cooper *et al.* (1974)	5		•		•			•	•	•		
Williams and Cavanagh (1983)**	8	•	•	•	•		•	•	•	•		
Studies that did not use swing events												
Barrentine *et al.* (1994)	0											
Neal (1998)	0											

* Carlsoo (1967) used eight swing events but only four events were clearly defined
** Williams and Cavanagh (1983) used eight swing events but did not report data at these events

Generally, the swing events used to quantify weight position has been at best inconsistent. For studies that analysed data statistically, between two and six swing events have been used. While two descriptive studies used eight swing events (Carlsoo, 1967; Williams and Cavanagh, 1983), limitations existed with both. Williams and Cavanagh (1983) provided only qualitative analysis of weight position at each event. Further, while Carlsoo (1967) did provide data, four of the eight swing events used were difficult to interpret from the definition given (refer to table 2.3).

The use of swing events is useful to make the data more interpretable and to allow for comparison between golfers (e.g. using time before BC does not allow for comparison of Fz% or CP% between subjects) but the reduction in data that exists in the literature has eliminated potentially important information. As noted in coaching and scientific literature, weight transfer changes direction before the top of backswing and moves rapidly forward in early downswing. This would indicate that late backswing and early downswing are potentially information-rich, a possibility supported by a significant result reported at FH by Robinson (1994) - the only study that has examined weight transfer at an event in either of these phases of the swing. As well, as discussed, Carlsoo (1967) and Cooper *et al.* (1974) reported movement towards the back foot in mid follow through, an event was not used in any of the studies performing statistical analyses on the data. This might also be a useful swing event to use. The use of more swing events to quantify weight transfer is addressed in this study.

2.1.5.4 The lack of examination of different swing styles between golfers

A possible limitation of previous studies examining weight transfer is the lack of examination of swing styles or movement strategies in the golf swing prior to performance analysis. A style or movement strategy is the performance of a skill in different ways to achieve the same aim. Two examples of obviously different styles in sport are single versus double handed backhand in tennis and the slide approach compared with the rotational approach in shotput. In golf, Meadows (2001) reported three different strategies for gripping the golf club: the interlocking grip, the overlapping grip and the baseball grip.

Less obvious is the existence of styles in the golf swing itself. However, different swing styles of professional golfers are often discussed in the media and have been noted in the coaching literature. For example, Suttie (2006) reported that, based on his “long years of teaching” all golfers fitted best into one of four categories based on the golfer’s physical traits and how the golfer develops power: the upper body, the lower body, the hands and a swing style termed “Classic” by Suttie which displayed an “unusually high degree of rhythm, timing and balance”. Upper body players possess shorter than normal arms and developed power from the shoulder muscles. Lower body players are taller, athletic players with long legs from which power is developed. Hands players possess large hands and forearms from which power is developed. Classic players are smooth, effortless and balanced, as well as being athletic with long arms.

Different styles in the golf swing have been reported, albeit with no objective data, in the scientific literature. Nagano and Sawada (1977) reported two styles of swing arc based on observation of the trace of the club head viewed in the vertical plane parallel to the line of shot. Swing golfers exhibited an elliptical orbit from takeaway to the end of follow through while ‘hitters’ exhibited a horizontal or oblique type swing. This was the limit of the definition of the swing styles. Neal (1998) used a golf coach to identify two styles of swing: a left-to-right and a rotational style of swing, although no definition of the two styles was presented. However, objectively assessed differences in swing styles or strategies do not exist.

Only two golf studies examining weight transfer have considered styles, or movement strategies, in their research design. In Neal (1998), a golf coach subjectively defined two styles in low handicap golfers; a 'rotational' style and a 'left to right' style. Koslow (1994) used a Sporttech swing force plate data to identify different weight transfer styles in novice golfers and reported three statistically different styles:

1. Normal (balanced at address, towards the back foot in backswing and towards the front foot at ball contact. Koslow reported this sequence as the "coach defined method of weight transfer").
2. Reverse pivot (balanced at address, on front foot at the top of backswing and on the back foot at ball contact).
3. Incomplete (balanced at address, back foot at backswing and back foot at ball contact).

Koslow (1994) reported that 84% of novice golfers did not exhibit the Normal pattern, producing instead a Reverse pivot or an Incomplete weight transfer. It is important to note that although different weight transfer patterns existed in the Koslow study, two of the three styles were considered errors by the researchers rather than useful techniques. This is an important distinction because different movement strategies within a skill can be valid methods of performance, or can represent errors in movement that reduce performance. That is, a particular style can be effective or it can be ineffective.

Neal (1998) compared two coach-defined and coach-identified styles among 14 low handicap golfers. These groups were defined as 'left-to-right' and 'rotational' but the criterion for each style was not reported. As low handicap golfers (i.e. highly skilled) were used, both of these styles might be considered to be valid techniques and not technical errors, although this was not stated by Neal. Differences were found between these groups in the ratio between the CP range perpendicular to the line of hit (CPx) and CP range parallel to the line of hit (CPy) between the groups (Left to right: 0.29, Rotational: 0.39, $p < 0.05$). Neal reported that the 'left-to-right' style produced greater CPy movement, or movement between the left and right foot (hence the 'left-to-right' term for this style) and less CPx movement compared to the rotational style ($p < 0.05$). As well, the 'left-to-right' group reached maximum forward position of CPy later in the swing (expressed as a percentage of the time between top of backswing and ball contact - Left to right: 99%, Rotational: 87%, $p < 0.05$). Neal (1998) did not report a description of the styles and, statistical analysis results were limited to alpha levels of $p < 0.05$ only, limiting further discussion.

If different styles existed in the samples tested in previous weight transfer research, statistical errors would have been made. Styles such as the two described by Neal (1998) occurring within the same dataset could generate type 1 or 2 errors in statistical analyses. Type 1 errors might exist in data due to the existence of two distinct groups (i.e. two distinct styles). Two groups (styles) can produce a large effect in regression statistics (e.g. figure 2.9 a) due only to the relatively large difference between the groups, with no relationship evident within the groups themselves. It could also produce type 2 errors in the case of comparison of handicap groups by increasing intra-group variance and reducing statistical power. Bates (1996) stated that if different movement strategies exit, then statistical power will be lowered and there is a greater likelihood of falsely supporting the null hypothesis in group comparison statistical procedures. Type 2 errors might also exist in regression statistics where no significant correlation exists for the whole group while significant correlations do exist within each group (e.g. figure 2.9 b). The lack of attention to possible style differences (or different movement strategies) may be a reason why many studies failed to find significant associations between weight transfer and performance, or why significant results were found in some studies.

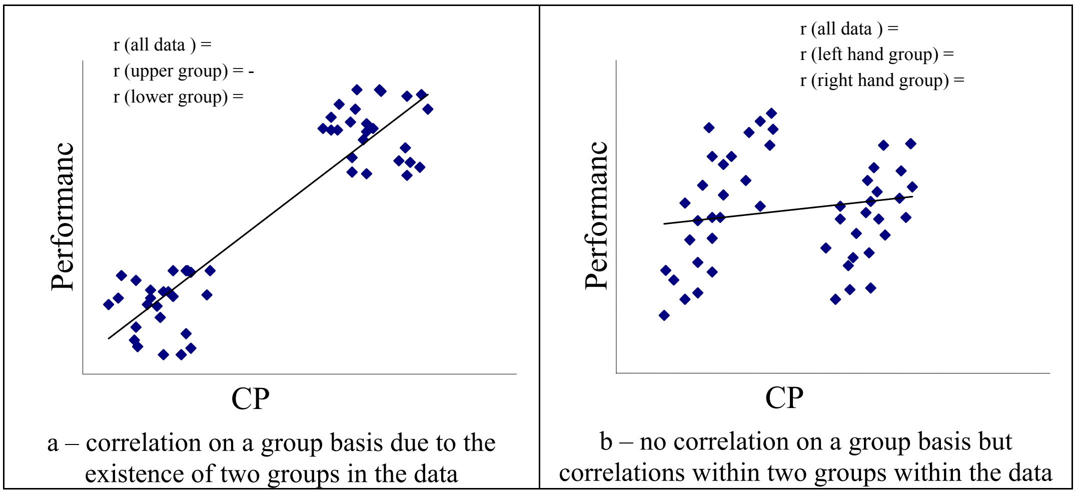


Figure 2.9: Examples of statistical errors that can exist where different styles or movement strategies exist within the dataset.

2.1.5.5 The lack of examination of individual-based analysis

Individual-based analysis, or single subject design, is the evaluation of a problem within a single subject. In its strictest sense, it involves examining one subject only and performing analysis within that subject (Reboussin and Morgan, 1996). However, much of the more recent work has been of a multiple single subject design, where a number of subjects are analysed on an individual basis and the research question is answered in terms of the number of significant results across all individuals (e.g. Ball *et al.*, 2003b; Dufek *et al.*, 1995). Individual-based examination has taken the form of correlation analysis to examine associations between technical and performance variables in shooting (Ball *et al.*, 2003a, 2003b), differences between means for different shoe conditions in running (Caster and Bates, 1995) and multiple regression predicting dependent variables in running (Dufek *et al.*, 1995).

A number of arguments exist for the use of individual-based analysis.

First, it can provide important information to the golfer that might be masked in group-based analyses. Significant performance factors have been found in individual analyses that were not evident in group-based analyses of pistol (Ball *et al.*, 2003a) and rifle shooters (Ball *et al.*, 2003b). Based on these findings, Ball *et al.* encouraged the use of both group and individual-based analyses to extract all relevant information from an analysis. Both Dufek *et al.* (1995) and Bates *et al.* (1979) also identified different performance characteristics among runners that were masked by a group-based statistical approach.

Second, individual-based analysis can avoid the statistical errors that will be produced if different movement strategies, or styles of movement, are used by different subjects in the group being tested. Bates *et al.* (2004) stated that individual performance strategies can lead to increased inter-subject variability that will reduce statistical power. This can then lead to false support for the null hypothesis.

Third, a number of authors have expressed concern with the group-based approach in some areas of clinical biomechanics research as most clinical applications are individual-based. For example, Scholhorn *et al.* (2002) noted that while the scientific literature is largely group-based, it does not reflect clinical practice in gait analysis, which is individual-based. Barlow and Hersen (1984) stated that many researchers were dissatisfied with the inability to find strong associations to explain behaviour that were being observed in the applied setting because of the use of group-based statistics. As golf coaching is largely individual-based, this argument would seem to be applicable to golf research.

While descriptions of individual golfers have been reported in the literature, there has been no individual-based statistical analysis of weight transfer. Carlsoo (1967) examined 300 swings of an elite golfer and used the data to describe the mechanics of the golf swing. Both Williams and Cavanagh (1983) and Richards *et al.* (1985) noted that the CP or COV variability was low within subjects tested. Further, Richards *et al.* (1985) stated that within-subject variability was similar for all ability levels tested (low and high handicap golfers). However, no study presented data to support these comments, nor was there any further analysis performed.

As no extensive individual-based analysis has been performed in golf, two examples from the literature have been discussed to illustrate the use of individual-based analysis.

In an elite sport example, Ball *et al.* (2003a) examined rifle shooters on a group and an individual basis. Six elite shooters performed 20 shots at a target under simulated competition conditions. Body sway and aim point fluctuation measures were correlated with performance to identify important factors in rifle shooting. Ball *et al.* reported finding no significant associations between body sway and performance on a group basis, although Ball *et al.* noted that statistical power was low in this analysis. However, all shooters returned significant correlations and regressions when the relationships were examined on an individual basis. Further, regression analysis showed individual-specific relationships for different individuals. Ball *et al.* stated that individual-based analysis is the most appropriate for the individual in terms of skill development and serves as a useful adjunct for group-based analysis. Ball *et al.* (2003b) also found differences between group-based and individual-based analyses for pistol shooters.

In a movement strategy example, Caster and Bates (1995) assessed the response strategies during landing for four subjects. Different masses were added to ankles to alter landing forces and the method of landing, defined by the researchers as either a mechanical or neuromuscular strategy, was examined. ANOVA and stepwise multiple regression was performed on ground reaction forces and EMG data on a group and individual basis. Caster and Bates reported that, while no group differences were found, 16 of 32 conditions were significant when examined on an individual basis. Further, multiple regression analysis showed individual-specific relationships in terms of the independent variables entered and the proportion of variance accounted for. As such, individual analysis allowed for the identification of strategies while group-based analysis did not. Based on the differences between the group and individual results, Caster and Bates concluded that it was inappropriate to use group-based analysis in this instance as it did not allow for the identification of these strategies. Similarly, Bates and Stergiou (1996) concluded that subjects can and do respond differently to the same perturbation and these differential responses can compromise group analysis results.

These five limitations of previous research in weight transfer in the golf swing – use of Handicap to assess swing quality, the small number of trials per golfer and swing events used, the lack of accounting for swing styles and the absence of individual-based analysis - will be addressed in this study.

2.2 CLUSTER ANALYSIS

Classification of objects that are thought to be similar to each other is one of the most inherent tasks for humans (Hair *et al.*, 1995). A quantitative method that has been widely used to classify objects is cluster analysis (e.g. In Medicine: Wang *et al.*, 2002; Psychology: Hodge and Petlichkoff, 2000; Biomechanics: Vardaxis *et al.*, 1998). Cluster analysis is a statistical process that uses multivariate techniques to group objects based on their characteristics. The aim of cluster analysis is to maximize the homogeneity of objects within a group, while maximizing heterogeneity between the clusters (Hair *et al.*, 1995).

Cluster analysis has been used in a small number of biomechanical studies and for a number of reasons. Wilson and Howard (1983) and Forwood *et al.* (1985) used cluster analysis to identify different movement patterns in a sporting skill with the aim to provide a condensed description of the skill. Grabe and Widule (1988) used cluster analysis to establish if the different skill level classifications could be objectively determined based on kinematic variables. Vardaxis *et al.* (1998) used cluster analysis to establish if different gait patterns existed in able-bodied males. In each study, the researchers report that cluster analysis proved useful for the particular application.

However, in spite of its potential usefulness, cluster analysis has not been widely used in biomechanics. One reason why cluster analysis has largely been avoided may be concerns of researchers with some of the procedural problems that still exist in the cluster analysis process. The major problems are how to determine the number of clusters that exist in the data (e.g. Hair *et al.*, 1995) and how to validate the cluster solution (e.g. Milligan, 1996). This is further complicated by the number of different clustering strategies (i.e. how cases are clustered together) as well as the many different methods for measuring the degree of similarity between cases (i.e. how the distance between cases is measured). These factors will be discussed in the next sections.

Note that this is a large field with a large body of literature pertaining to it. It is not the intention of this review to be comprehensive in all aspects of cluster analysis. Rather, the points of contention that can affect the analysis are discussed with an emphasis on biomechanical research using clusters. The researcher refers the interested reader to Hair *et al.* (1995) for a good overview of operational information and general issues.

2.2.1 How many clusters?

Cluster analysis will always return a solution. As such, the choice of the number of clusters in the final solution is extremely important (Hair *et al.*, 1995). Figure 2.10 shows two plots with the same data but divided into two or three clusters. Each might be considered a reasonable division of the data into groups (as would only one cluster; i.e. the group has no clusters within it). As the one, two and three cluster solution will each lead to a different conclusion, it is important to thoroughly assess which solution is most appropriate. The numerous methods used to address this issue include the use of the agglomerative schedule (also called ‘stepwise method’) and dendrogram (e.g. Forwood *et al.*, 1985), various statistical procedures (e.g. the R-Ratio: Vardaxis *et al.*, 1998), and the use of theoretical assessment (e.g. minimum number of meaningful clusters; Wilson and Howard, 1983).

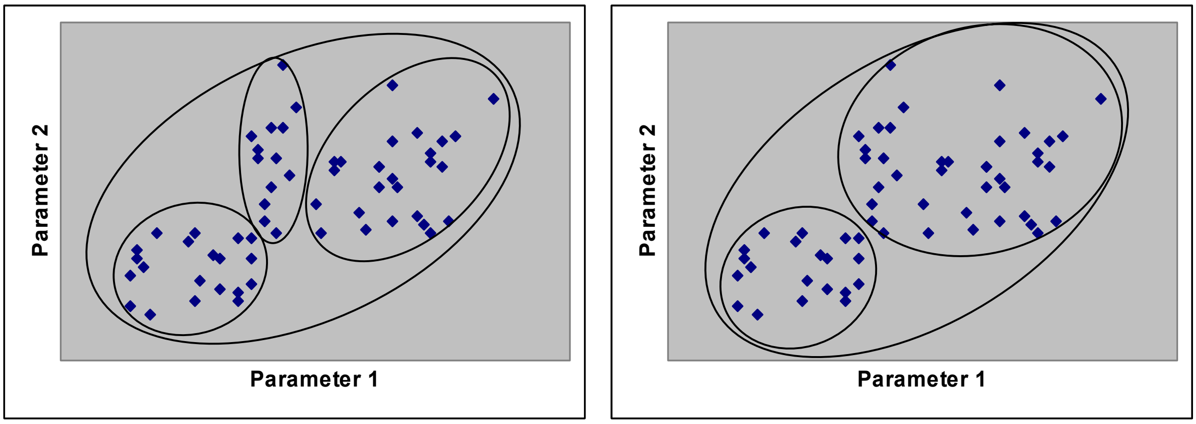


Figure 2.10: Example of one, two and three cluster solutions for a two-parameter dataset. Possible clusters indicated within each ellipse.

2.2.1.1 Agglomerative schedule and dendrogram

The use of the agglomerative schedule is noted in the SPSS 10.0 manual (1997) as the most appropriate method for choosing the number of clusters when performing hierarchical cluster analysis. Briefly, cases are clustered together progressively, starting with N clusters (i.e. each case is a cluster with $N = 1$ cases in each cluster) and then sequentially combining the two nearest clusters until there is only one cluster (i.e. all N cases in the one cluster). This process can also be performed in reverse, starting with one cluster and progressing to N clusters. The distance between clusters is calculated at each step and reported in the agglomerative schedule. Thorndike (1953) suggested that the researcher should look for a sharp step in this data, which indicates that clusters are more widely separated and can indicate an optimal solution. Gower (1975) noted that unless a jump in this stepwise method exists, there is no argument for defining any more than one cluster.

The dendrogram (figure 2.11) provides a visual representation of this process and allows for identification of likely clusters as well as possible outliers in the data. The cases on the vertical axis represent the objects being clustered together. The horizontal axis scale represents the distance between clusters for consecutive combinations of cases. The lines indicate when cases combine together. Examining these lines for figure 2.11, the horizontal distance between consecutive joins on the right hand side (i.e. the last two joins) are relatively larger than those near the left hand side of the dendrogram. These large jumps are associated with the two and three cluster solutions. The one that is largest (the 2-cluster solution, or the last join) would be indicated as optimal by the dendrogram.

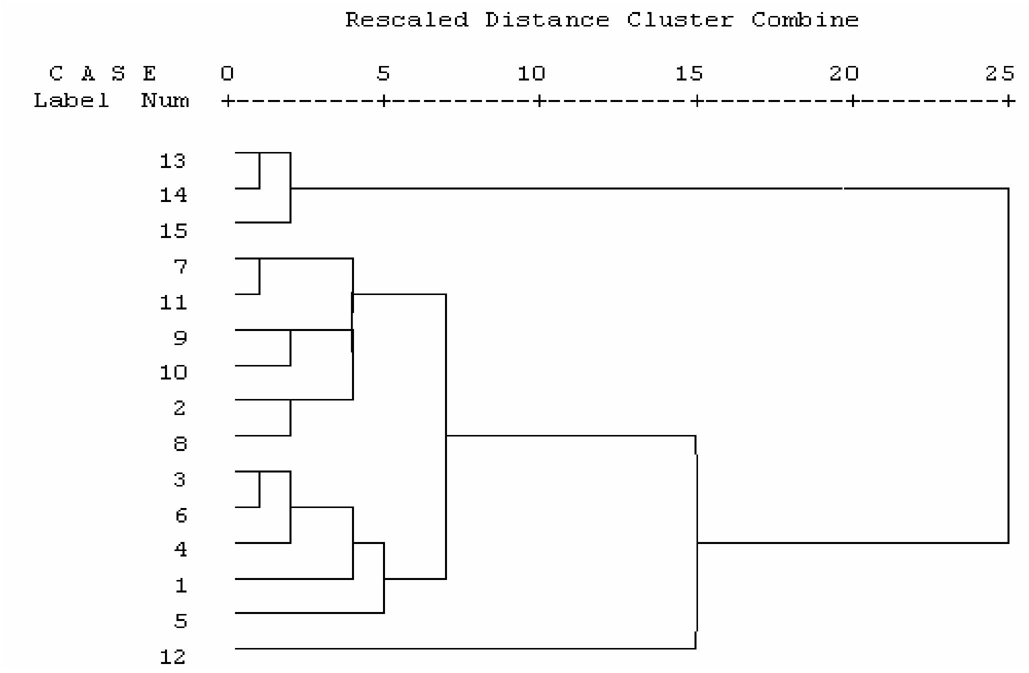


Figure 2.11: Example dendrogram with 15 cases (left hand side) clustered progressively into 1 cluster (right hand side). The scale on the horizontal axis indicates the separation between clusters.

However, the use of the agglomerative schedule and dendrogram method to determine the number of clusters has been criticised. Everitt (1979) noted that the choice of what jump equates to the 'sharp step' referred to by Thorndike (1957) is likely to be subjective. Further, Everitt (1974) found that a large step is a necessary but not a sufficient condition that clear-cut clusters exist in the data. Milligan and Cooper (1985) described the use of the agglomerative schedule to decide on the number of clusters (termed the "stepwise method") as "mediocre". Interestingly, Hair *et al.* (1995) suggested that the agglomerative schedule method returned reasonable results, even though Hair *et al.* refer to the Milligan and Cooper (1985) article as the source of this interpretation.

2.2.1.2 Statistical methods for determining cluster solutions

The number of clusters in a dataset has also been determined by statistical procedures. These methods are often called 'stopping rules' as they indicate where in a hierarchical clustering process the clustering, or joining/separating of groups, should stop. For example, the C-Index (equation 2.3; Hubert and Levin, 1976) calculates the distance between each case within each cluster in a solution and then sums them to form a within-cluster distance (D, equation 2.3). The optimal solution, or number of clusters, is associated with the largest C-Index score produced across all cluster solutions (from 2 to *N*, Hubert and Levin).

$$\text{C-Index} = \frac{\text{Sum (D - all clusters) - Minimum (D)}}{\text{Maximum (D) - Minimum (D)}}$$

Where D = sum of the distances between all cases within a cluster

Equation 2.3

Numerous statistical methods have been used in the literature (e.g. C-Index: Hubert and Levin, 1976; Goodman-Kruskal's Gamma: Baker and Hubert, 1975; Point Biserial Correlation: Milligan, 1981). Milligan and Cooper (1985) evaluated the performance of 29 of these statistical measures for determining the number of clusters as well as the agglomerative schedule method. Using 432 Monte Carlo datasets with a known number of clusters in the data, the 30 stopping rules were applied to each dataset. Briefly, the Monte Carlo data set is a pseudo-random collection of numbers based around forming a known number of clusters. The stopping rules were then ranked based on the number of correct solutions that were returned. Milligan and Cooper identified the better stopping rules and suggested that the use of a combination of the better statistical methods is the most appropriate procedure in establishing the number of clusters in a dataset, a point stressed again in later work by Milligan (1996).

The use of statistical procedures to identify the number of clusters in a dataset has been criticised or cautioned in the literature. Hair *et al.* (1995) suggested that while statistical techniques might provide a more objective measure compared to the agglomerative schedule and dendrogram, most statistical methods are overly complex for the benefit they provide. Milligan and Cooper (1985) and Everitt (1979) noted that many of these statistical procedures are data specific and may not work well on data that they were not developed for, although neither indicated any data to support the comment. As well, Milligan and Cooper (1985) noted that different stopping rules exhibited different desirable properties. For instance the Calinski and Harabasz index returned the most number of correct solutions, while the Point Biserial correlation had the lowest error rate for determining too many clusters. Further, with the many possible statistical methods available requiring a choice, the subjectivity of which to use can offset some of the objectivity they provide.

2.2.1.3 Theoretical considerations

Hair *et al.* (1995) suggested that the final decision on how many clusters are formed should have theoretical considerations to support the empirical evidence provided by the various statistical and graphical techniques. As well, Milligan (1996) noted that if the clusters are not interpretable, then the analysis is of no use. As such, theoretical considerations must form a part of any decision on the number of clusters. Surprisingly, while many researchers have developed statistical methods for determining the number of clusters, little research has dealt with theoretical issues that are associated with this process, beyond noting that a theoretical basis is important.

2.2.1.4 Method of choosing the number of clusters in biomechanical research

The biomechanical studies that have used cluster analysis have employed different combinations of these factors to decide on the number of clusters in their data. Both Wilson and Howard (1983) and Forwood *et al.* (1985) used a combination of visual inspection of the dendrogram and the theoretical consideration of finding the minimum number of movement action patterns that could describe the backstroke start (Wilson and Howard, 1983) and the front handspring to front somersault in gymnastics (Forwood *et al.*, 1985). The basis for selection of the optimal solution from the dendrogram was not detailed in either article. Vardaxis *et al.* (1998) used the dendrogram and the R-ratio. The R-ratio is a measure of the reduction of the within-cluster variability, with a large R-value associated with a large reduction in within-group variability and indicating that clusters are homogenous. The largest R-ratio across different cluster solutions was considered to be optimal. The limitation of the R-ratio is that there is no between-cluster evaluation and so should have been used with other indicators that provide between-cluster information. Milligan and Cooper (1985) did not test this method.

Grabe and Widule (1988) used an unusual approach to choosing the number of clusters that existed in kinematic data of the jerk in weightlifting. A two-way coincidence table was calculated which compared ordered subjects (weightlifters) with ordered variables (29 kinetic and kinematic variables associated with the Jerk movement in the Olympic Clean and Jerk lift), with meaningful clusters identified using a variable (rather than fixed as the methods discussed to date have been) stopping rule as outlined by Boesch (1977). The Boesch article could not be obtained for inclusion in this review and as the details were not thorough enough in the Grabe and Widule (1988) article, further discussion of this stopping rule is not possible.

2.2.2 Validation of clusters

Validation of cluster solutions is an important part of cluster analysis. Aldenderfer and Blashfield (1984) noted that while validation techniques are not well understood and should be used with caution, it is essential to validate the solution obtained from a cluster analysis. Hair *et al.* (1995) suggested that without validation, the cluster solution found has limited generalisability and will have little use beyond a description of the data on clustering variables.

2.2.2.1 Methods of validation

Milligan (1996) reported that validation methods are broadly divided into external and internal methods. These have been reported separately in the next sections.

2.2.2.1.1 EXTERNAL METHODS

External methods of validation are based on variables not used in the clustering process. Milligan (1996) reported that two types of external validation are available; (1) the use of standard parametric procedures on a variable not used in the clustering process to test for differences between clusters and (2) the use of an independently obtained partition.

2.2.2.1.1.1 Parametric tests on an external variable

The use of an external variable has been described as among the better methods of validating a cluster solution (Aldenderfer and Blashfield, 1984). Briefly, a variable, which has a theoretical basis for being different between groups or for defining the clustering solution, is identified. Parametric tests such as ANOVA are then performed on this variable to compare between groups. A significant result supports the validity of the cluster solution. For example, the coach defined ‘left-to-right’ or ‘rotational’ swing style in the Neal (1998) study could have been used as an external variable for cluster analysis on swing characteristics to determine if two styles exist.

However, while the use of external variables to validate a cluster solution has been supported, a number of difficulties are associated with it. Aldenderfer and Blashfield

(1984) reported that the identification of an external variable may not be possible due to the research being exploratory, or the theoretical basis behind the cluster analysis may not be refined sufficiently to determine what is relevant to the intended classification. As well, Milligan (1996) noted that researchers find it difficult to omit a variable from the clustering process with the aim of using that variable to externally validate the cluster solution. Further, even if an external variable is used to validate a cluster solution, the validity of the external variable itself in an applied research setting can be questioned. Milligan (1981) noted that if the researcher fails to find a significant difference between the groups on an external variable, it is not necessarily clear as to whether this indicates a lack of cluster structure, or just that external variable used to validate the clusters is invalid itself.

2.2.2.1.1.2 Independently obtained partition

Milligan (1996) detailed another external validation method in which a partition (clustering pattern) is specified a-priori or obtained from clustering of another dataset. The cluster solution found in the data is then compared with the specified solution, with a high level of agreement equating to a more strongly validated solution. Milligan reported that a number of indices have been proposed such as the Rand index and the Jaccard statistic (e.g. Rohlf, 1974). However, in applied research, the true cluster structure, which is required for these indices, is generally not known a-priori (Milligan, 1981). None of the applied studies reviewed by this researcher used these external criterion indices.

2.2.2.1.2 INTERNAL METHODS

Internal methods of validation use the variables that have already been used in the clustering process. These methods attempt to represent the goodness of fit between the input data and the resulting cluster solution (Milligan, 1996). Internal methods are closely linked to the choice of the number of clusters and many of the same statistical techniques are used.

2.2.2.1.2.1 Statistical methods

The SPSS 10.0 manual (1997) suggests that the best way to validate a cluster solution is conducting a discriminant analysis on the clustered data. In this method, if the discriminant analysis indicates that the groups (clusters) are significantly different, then the solution is validated. However, the use of discriminant analysis, ANOVA and MANOVA has been strongly criticized (e.g. Milligan, 1996; Aldenderfer and Blashfield, 1984). Milligan (1996) argued that the clustering process separates cases into groups that minimize overlap, and so techniques such as discriminant analysis, ANOVA and MANOVA will always show good results when the cluster groups are compared. Interestingly, while Hair *et al.* (1995) also noted this method is inappropriate, the researchers use a one-way ANOVA to compare cluster groups in an example presented in their work, although it was not used to validate the solution. It would seem that the use of these tests on clustering variables still offers useful information to the researcher in establishing which variables differ between groups, but this information does not validate the cluster solution on its own.

Milligan (1981) compared 30 methods for internal validation, using 108 Monte Carlo datasets with known cluster solutions and using two external criteria measures (Rand and Jaccard statistic). Applying each statistical technique to the clusters, the study ranked the 30 internal methods based on how close the result was to the known clustering solution as indicated by the external criteria measures. Milligan identified a group of six 'strong' methods that could form the basis of a validation procedure in applied research. Milligan (1996) noted that this testing was conducted on artificial data and may not hold for real world data. However, it does offer strong objective data for the relative effectiveness of different methods that can be used by the researcher to guide the choice of a validation method.

2.2.2.1.2.2 Replication

Replication has been reported as a possible method for validating cluster analysis (e.g. McIntyre and Blashfield, 1980; Morey *et al.*, 1983). Replication refers to the process of repeating the cluster analysis on a randomly drawn subset of the original data. If a cluster is robust (i.e. if its characteristics remain despite the use of different sub-sets from the sample population) the researcher has some evidence to support the solution's existence. Milligan (1996) noted that replication is analogous to the cross-validation procedure in regression analysis. Replication was used by Hodge and Petlichkoff (2000), for example, reanalyzing a randomly selected subset of two thirds of the original dataset. The researchers reported that 94% of the subset subjects maintained the same cluster membership as the original analysis concluded that the solution was robust based on these results.

Milligan (1996) reported a slightly different replication method. Two samples of data are obtained (usually by randomly dividing the initial dataset). Cluster analysis is performed on the first dataset and means for each cluster are calculated. Using these means, each case from the second dataset is allocated to the nearest cluster, and the cluster that each case is allocated to is noted for later comparison. Then the second dataset is cluster analysed. The two cluster solutions for the second dataset (i.e. from cluster means from dataset one and from cluster analysis of dataset two) are then compared. The level of agreement between the two cluster solutions reflects the stability of the cluster solution. Breckenridge (1989), as reported by Milligan (1996), found this method useful for validating clusters in work with Monte Carlo datasets.

However, Aldenderfer and Blashfield (1984) criticised the replication method of validation. They suggested that finding a similar cluster structure using replication is a check of the internal consistency of the result. While the failure of a cluster solution to be replicated is reason for rejecting the solution, or the existence of an individual cluster, successful replication does not guarantee validity of a solution. Unfortunately, Aldenderfer and Blashfield did not expand on this issue. No other author has expressed concern with replication as a validation method. As well, replication is a method being used more extensively in association with other statistics (e.g. regression, Pedhazur, 1997). Re-sampling methods such as bootstrapping and jack-knifing (Zhu, 1997) are examples of replication analyses that are being used more extensively to provide confidence limits and validation to solutions using other statistics (e.g. Ball *et al.*, 2003a). It should also be noted that no single measure (replication, statistical measures, theoretical assessment) completely validates a solution but all should be included to build up support for the cluster solution validity.

2.2.2.1.2.3 Use of more than one clustering algorithm

The use of more than one method or measure of cluster calculation has been proposed as a useful method of validation of cluster analysis. There are a number of methods within the cluster process for measuring the distance between cases and between clusters. For example, the 'between-groups' method clusters cases that maximize the distance between clusters at each step of the hierarchical process while the 'within-groups' method simply clusters the two nearest cases or clusters at each step. There are also a number of measures used to define how the distance between cases/clusters is quantified and include Euclidean distance, squared Euclidean distance (referred to as measures of dissimilarity) and Pearson's correlation (referred to as a measure of similarity).

Hair *et al.* (1995) suggested that re-analyzing a cluster solution using non-hierarchical techniques with random selection of starting seeds is a way to test the robustness of the cluster solution and validate results. In the example Hair *et al.* presented, the initial solution, calculated using hierarchical techniques, was found. Then, using a non-hierarchical clustering process and using k random seeds (i.e. randomly selected cluster centroids or means) and where k equalled the number of clusters chosen from the hierarchical process, clustering was performed on the data again. Hair *et al.* reported that this non-hierarchical cluster analysis obtained the same clusters as the hierarchical procedure and, based on this finding, concluded that the solution is robust and valid. However, Milligan (1996) reported that, while non-hierarchical procedures with known seed points are better at obtaining correct cluster numbers than hierarchical procedures, if random seeds are used, clustering is poor. This being the case, the method suggested by Hair *et al.* (1995) would seem to be inappropriate. Interestingly, Hair *et al.* also noted that using a non-hierarchical method with random seeds leads to poor clustering solutions.

Kos and Psenick (2000) suggested that added validity is provided to a cluster solution if the clusters appear using different methods for measuring the distance between cases and clusters. Kos and Psenick clustered a dataset using both the between-group method and within-group method, and suggested that for a cluster to be considered valid, it must appear in both analyses. Hair *et al.* (1995) also suggested that the use of more than one clustering method would be an appropriate way to validate a cluster solution, although the researchers do not use it in the example they provide. However, as noted by Hair *et al.* (1995) and Milligan (1996), the choice of method of determining clusters should have a theoretical basis. The use of more than one method suggested by Kos and Pesnik (2000) may depart from this theoretical basis in many applied research applications, making the results of such comparisons invalid themselves.

2.2.2.1.2.4 Monte Carlo datasets

Aldenderfer and Blashfield (1984) suggested Monte Carlo data sets might be a useful method of validation of a cluster solution. In this case, the Monte Carlo data set is generated so that its characteristics are the same as the characteristics of the original data set (such as means, standard deviations etc.) but with no pre-defined clusters. Both the original and Monte Carlo data sets are then cluster analysed. Aldenderfer and Blashfield suggested that the next step could involve performing one-way ANOVA on each of the parameters between clusters for the original data set. Similarly, one-way ANOVA between clusters from the Monte Carlo data set is also performed. If the difference in the F-ratios between the original and Monte Carlo data sets is large, then it might be considered that the cluster solution is sufficiently removed from a random result to be considered valid. Conversely, if F-ratios were similar then little support exists for the cluster groupings being valid, and more likely exist due to chance. Aldenderfer and Blashfield noted that this method had not been widely used (by 1984) and this researcher found little use of it in the literature since this time. Milligan used this technique in the series of studies evaluating different cluster methodologies but not to validate a real data solution (e.g. Milligan, 1981; Milligan and Cooper, 1985).

2.2.2.1.2.5 Cophenetic correlation

Aldenderfer and Blashfield (1984) also discussed the use of cophenetic correlation to validate a cluster solution. Briefly, this method examines the dendrogram to see how well it represents the pattern among the clustered cases. An implied similarity matrix is developed based on when cases were clustered together. For example, similar cases will cluster together early in the process and these cases will incur a small value. Conversely, dissimilar cases will cluster together late and will incur a large value. This matrix is compared with the original matrix obtained from the Euclidean distance between cases. The cophenetic correlation is the correlation between values in the original and implied matrix, with a larger value indicating a better clustering of the data.

Aldenderfer and Blashfield (1984) were critical of the use of cophenetic correlation, suggesting that the assumption of normality (required for correlation) is usually violated and so the correlation coefficient is not an optimal estimator of the degree of similarity between the two matrices. This might not be a valid criticism given numerous authors suggested correlation is relatively robust to violations of non-normality (e.g. Tabachnick and Fidell, 1996) and non-parametric tests might avoid this problem (Aldenderfer and Blashfield did not discuss this possibility). However, other limitations exist. The technique can only be used on data that has been clustered using the hierarchical method. As well, Aldenderfer and Blashfield reported that the two matrices contain different amounts of data and so contain considerably different information. Further, Holgersson (1978), as reported by Aldenderfer and Blashfield (1984), found the cophenetic correlation to be a generally misleading indicator of cluster quality based on assessment using Monte Carlo datasets.

2.2.2.2 Validation in biomechanical research

The use of cluster validation in biomechanical research has been limited. Wilson and Howard (1983), Forwood *et al.* (1985) and Vardaxis *et al.* (1998) used one-way ANOVAs on the variables used in the clustering process to assess if clusters were significantly different, a method criticized by a number of researchers (e.g. Aldenderfer and Blashfield, 1984; Milligan, 1996). Vardaxis *et al.* (1998) also tested the robustness of the cluster solution by randomly eliminating subjects, and reported that the solution remained unchanged with up to 15% of the total sample size eliminated. This is something similar to the replication methods reported in section 2.2.2.1.2.2. Vardaxis *et al.* (1998) also reported using the MGHL procedure of SYSTAT, which tests the null hypothesis that clusters are equal. SYSTAT was not available to this researcher and so this method could not be evaluated. Grabe and Widule (1988) did not validate their cluster solution.

2.2.3 Hierarchical or non-hierarchical methods

Another factor in the cluster process is the use of hierarchical or non-hierarchical methods for forming clusters. Hierarchical clustering techniques effectively construct a tree-like structure, with progressive clustering (agglomerative) or un-clustering (divisive) of cases. The agglomerative method is evident in the dendrogram in figure 2.11 (section 2.2.1.1). Non-hierarchical procedures assign objects into clusters, of which the number and the seeds (starting centroids or group means) are defined a-priori by the researcher.

Milligan (1980) has suggested that non-hierarchical procedures are more robust than hierarchical procedures in extracting the true cluster solution from the data, due to the elimination of nesting. Nesting occurs in the hierarchical procedure of clustering where a case is allocated to a cluster in an early step but in the final solution may be more appropriately allocated to another cluster. Non-hierarchical methods reduce the chance of nesting as each case is evaluated against a constant set of group means (i.e. the seeds specified by the researcher), compared with the constantly changing centroids in the hierarchical process. Hair *et al.* (1995) also supported the use of non-hierarchical methods, as they reduce the effects of outliers and the distance measure used (e.g. between-group or within-group measures). Hair *et al.* also noted that non-hierarchical methods reduce the effect of irrelevant variables (i.e. variables irrelevant to the underlying cluster structure), which can adversely affect hierarchical analysis by producing a different cluster solution. However, both researchers noted that this is only the case if the seeds, or starting points which non-hierarchical procedures require, are chosen carefully and are not random. Unfortunately, as noted by Milligan (1985), the identification of these seeds is often not possible as there is not enough previous data for the researcher to decide on an appropriate cluster seed. Milligan suggested a combined approach, with hierarchical methods used first to establish seeds (cluster means) and then using these seeds in a non-hierarchical cluster analysis to 'fine tune' the solution.

2.2.3.1 Hierarchical or non-hierarchical methods in biomechanical research

All biomechanical research using cluster analysis has used the hierarchical method. This may be a limitation of these studies. However, it may not have been feasible for the researchers to perform non-hierarchical analyses due to the lack of non-hierarchical software resources. Two of these studies used the Pearson's correlation as the measure of similarity (Grabe and Widule, 1988; Vardaxis *et al.*, 1998). The similarity measure was not reported in the remaining two studies (Wilson and Howard, 1983; Forwood *et al.*, 1985). This researcher could not find statistical software that offered a non-hierarchical process using the correlation method. SPSS, for example, offers only the Euclidean distance measure for use in the non-hierarchical cluster process. Milligan (1985) noted that the availability of appropriate analysis software is a major consideration in cluster analysis due to the calculation demands of assessing a large number of potential combinations. Wilson and Howard (1983) also noted the problems of computer power in their analysis.

2.2.4 Summary

In summary, there are a number of procedural issues in cluster analysis and this has been highlighted by the lack of consistent methodologies used. The choice of the number of clusters in the final solution must be considered carefully. Validation is an important part of cluster analysis and a sound knowledge of the processes is required. Finally, the choice of method for clustering needs to be considered.

Based on the literature, a strong approach would be to initially analyse the data using a hierarchical cluster analysis to establish cluster seeds and then reanalyze using non-hierarchical cluster analysis. The agglomerative schedule, dendrogram and possibly other techniques should be examined to establish if clusters exist in the data and that the data does not simply describe a continuum. Within the hierarchical and non-hierarchical processes, the choice of distance (similarity/dissimilarity) measure and the clustering method (e.g. between-groups, within-group) should be based on theoretical issues associated with the particular application (type of data, the research question being answered). The choice of number of clusters should include two or more of the better statistical measures recommended by Milligan and Cooper (1985) as well as theoretical considerations for the number and type of clusters formed. Validation should be performed using a combination of statistical measures, replication and theoretical assessment. While not supported by some authors, the use of ANOVA or MANOVA has been used widely in the literature and may still provide useful validation information if used in combination with other validation measures.

CHAPTER 3

GENERAL AIMS OF THE THESIS

The general aims of this thesis are to:

1. Identify if styles exist in the golf swing.
2. Determine if weight transfer is important in the golf swing on a group basis.
3. Determine if weight transfer is important on an individual basis.

These aims are examined across three studies in this thesis.

Study 1 will determine if different weight transfer patterns exist among golfers using a cluster analysis methodology

Study 2 will determine if weight transfer parameters are related to performance on a group basis using regression analysis.

Study 3 will determine if weight transfer parameters are related to performance on an individual basis using curve-fitting analysis and a non-linear technique.

CHAPTER 4

STUDY 1

IDENTIFICATION OF WEIGHT TRANSFER STYLES

4.1 AIMS

4.1.1 General

1. To examine if different weight transfer styles exist in the golf swing.
2. To compare skill level, performance and weight transfer parameters between different weight transfer styles if they exist.
3. To assess the cluster analysis methodology.

4.1.2 Specific

1. To apply cluster analysis to centre of pressure parallel to the line of shot at eight swing events to identify different weight transfer styles if they exist.
2. To compare between different weight transfer styles (if weight transfer styles are found to exist) for:
    - Handicap
    - Club Velocity
    - Descriptive data (Age, Height, Mass)
    - Centre of pressure parameters
3. To evaluate the usefulness and validity of cluster analysis for use in weight transfer in the golf swing

4.2 METHODS

4.2.1 Subjects

Sixty-two golfers ranging in skill level from professional players to high handicappers, as well as recreational golfers, were used in this study. Recreational golfers were defined as golfers who played at least 5 games a year but did not have a handicap. Subjects were canvassed from different golf clubs in Australia and included both right and left-handed golfers. Handicap, age, height and mass of the group are reported in table 4.1. A breakdown of handicaps is presented in table 4.2.

Table 4.1: Subject data

	Handicap	Age (years)	Height (m)	Mass (kg)
Mean	11.1	34.1	1.81	81.7
SD	8.0	13.7	0.07	9.1
Range	Plus 2 - 28	15 - 63	1.65 – 1.98	63.2 – 104.6

Note: 'Plus 2' is a handicap given to highly skilled golfers: 2 shots are added to their score.

Table 4.2: Handicap details of golfers used in this study

	Professional and/or tour players	0-4	5-9	10-14	15-19	20 +	Recreational golfers
Number of golfers	5	9	12	14	6	10	6

4.2.2 Task

All golfers were requested to bring their own golf club (driver), golf shoes and golf glove (if normally worn when playing) to be used in testing. In requiring golfers to use their own equipment, it was expected that a swing that was more indicative of their typical swing would be produced. After familiarisation with the laboratory environment and adequate warm up, each subject performed 10 swings using their driver, hitting the golf ball into a net placed 3 m away. Subjects were instructed to perform their typical swing. No time restrictions were placed on the golfers between trials.

4.2.2.1 Number of trials

The decision to use 10 trials to establish mean swing performance for each golfer was based on the results of techniques adapted from Bates *et al.* (1983). To identify the number of trials required for the mean centre of pressure (CP) values to stabilize, the method outlined by Bates *et al.* (1983) was used.

1. Five golfers performed fifteen swings under test conditions (i.e. the same conditions and testing protocol as outlined through section 4.2.2 – 4.2.3).

2. For each parameter used in this study (refer to section 4.2.5.2, tables 4.7 and 4.8 for definitions), the mean, standard deviation (termed Overall SD) and one quarter of the standard deviation (termed Threshold SD) were calculated for each golfer.

3. The next stage involved taking the first and second swing for a golfer and calculating a mean from these two swings, then adding the third datapoint and recalculating the mean. This continued until all fifteen trials were included and fourteen means were calculated (termed Floating Mean).

4. The change in Floating Mean between stages in step 3 was calculated (thirteen values resulted).

5. The threshold for stability was defined as the first change in Floating Mean that fell below the one quarter SD (Threshold SD) calculated from all swings (Bates *et al.*, 1983). The swing number at which this occurred was determined for each parameter and for each golfer.

6. In addition to this technique presented by Bates *et al.*, this researcher also examined the change in mean for trials after the threshold was reached (up to 15 trials) to see if the mean continued to remain stable.

An example analysis is presented in table 4.3. Overall standard deviation for all fifteen trials was 1.25 $m.s^{-1}$ (Threshold SD = one quarter times this value = SD * 0.25 = 0.31 $m.s^{-1}$) for Club Velocity for the selected golfer. In the column headed Floating Mean, the results of step 3 are presented. The first value in this column (49.7 $m.s^{-1}$) is the mean of the first and second trial. The next value (49.5 $m.s^{-1}$) is the mean of trials one to three and so on. The column headed Change in Floating Mean is the difference between consecutive Floating Mean values.

Table 4.3: Example analysis to determine the number of trials for the mean to stabilise adapted from Bates *et al.* (1983).

	Club Velocity (m.s^{-1})	Floating Mean	Change in Floating Mean
All trials (*N* = 15)			
Overall SD	1.25		
Threshold SD (SD * 0.25)	0.31		
Individual trials			
1	48.9		
2	50.6	49.7	
3	49.2	49.5	0.19
4	49.4	49.5	0.02
5	49.4	49.5	0.01
6	47.5	49.2	0.33
7	46.9	48.8	0.32
8	46.4	48.5	0.31
9	46.9	48.4	0.18
10	49.7	48.5	0.14
11	48.1	48.5	0.04
12	47.8	48.4	0.06
13	46.9	48.3	0.11
14	48.1	48.3	0.02
15	47.2	48.2	0.07

Note: Shaded cells indicate first (trial 3) and last (trial 9) time the Change in Floating Mean drops below the threshold of one quarter SD.

The first Change in Floating Mean value was below the Threshold SD (i.e. trial 3 – 0.19 is less than 0.31). Based on Bates *et al.* (1983) criteria, this indicated that Club Velocity would require three trials to stabilise for this golfer. However, the Change in Floating Mean at trials six, seven and eight were greater than the Threshold SD value. This occurred for approximately half of the golfer-parameter combinations where the change in Floating Mean did not remain below the Threshold SD. Based on this finding, the technique used by Bates *et al.* was modified. The criterion of the first Change in Floating Mean that fell below the Threshold SD was changed to the last Change in Floating Mean that fell below the

Threshold SD within the 15 trials evaluated. This was termed 'Last Threshold Cross' to contrast with the 'First Threshold Cross' suggested by Bates *et al.*.

Mean, maximum and minimum trial numbers for each parameter are presented in table 4.4. For the reader's information, both thresholds (first and last threshold cross) have been presented.

Table 4.4: Number of trials required for parameter means to stabilize (*N* = 5 golfers). Parameters are fully defined in tables 4.7 and 4.8 section 4.2.5.2.

	First Threshold Cross		Last Threshold Cross	
	Mean	Maximum	Mean	Maximum
Club Velocity	3.2	4	5.3	9
CPy%TA	3.2	4	6.0	8
CPy%MB	4.0	5	6.8	9
CPy%LB	3.6	4	5.4	10
CPy%TB	3.2	4	5.6	10
CPy%ED	3.2	4	6.2	9
CPy%MD	3.0	3	5.5	8
CPy%BC	3.6	5	5.5	9
CPy%MF	3.2	4	3.3	4
VelCPyTA	3.6	5	6.0	8
VelCPyMB	3.4	4	6.0	8
VelCPyLB	5.0	6	6.0	8
VelCPyTB	4.4	6	7.5	10
VelCPyED	3.6	5	5.0	8
VelCPyMD	3.2	4	4.8	7
VelCPyBC	4.0	5	6.3	9
VelCPyMF	3.6	4	6.3	9
VMaxCPy	3.6	5	5.0	6
tVMaxCPy	3.6	4	4.4	6
MaxCPy%	3.6	5	5.8	9
tMaxCPy%	3.4	5	3.5	5
MinCPy%	4.2	6	5.0	8
tMinCPy%	4.0	5	5.0	9
CPyR	3.2	4	5.0	8
CPyR%	3.6	5	6.2	8

Note: Maximums are single values taken from any of the five golfers tested (i.e. not average maximum across all five golfers)

Last Threshold Cross values indicated between three and eight trials were required, on average, for means to stabilise. However, single maximums were as high as 10 trials for some parameter-golfer combinations. Based in these results it was decided to use ten trials to obtain a stable mean. Also of note in table 4.4 is that no parameter produced the same set of results for both thresholds.

4.2.3 Laboratory set-up

While performing each swing, golfers stood on two AMTI force plates (Advanced Mechanical Technologies, Inc, Massachusetts, USA), one under each foot. The force plates were covered with Pro-Turf synthetic grass, as used in golf driving ranges and in golf shops. The golf ball was placed on a rubber tee, which was part of a ProV golf swing analyser (Golftek Incorporated, Lewiston, Idaho). A number of tee heights were available to the golfer and each selected their preferred tee during warm up swings. A large net was placed approximately 3 m in front of the hitting area to catch the ball after impact. All players reported feeling comfortable with the set-up. The set-up is represented in figure 4.1.

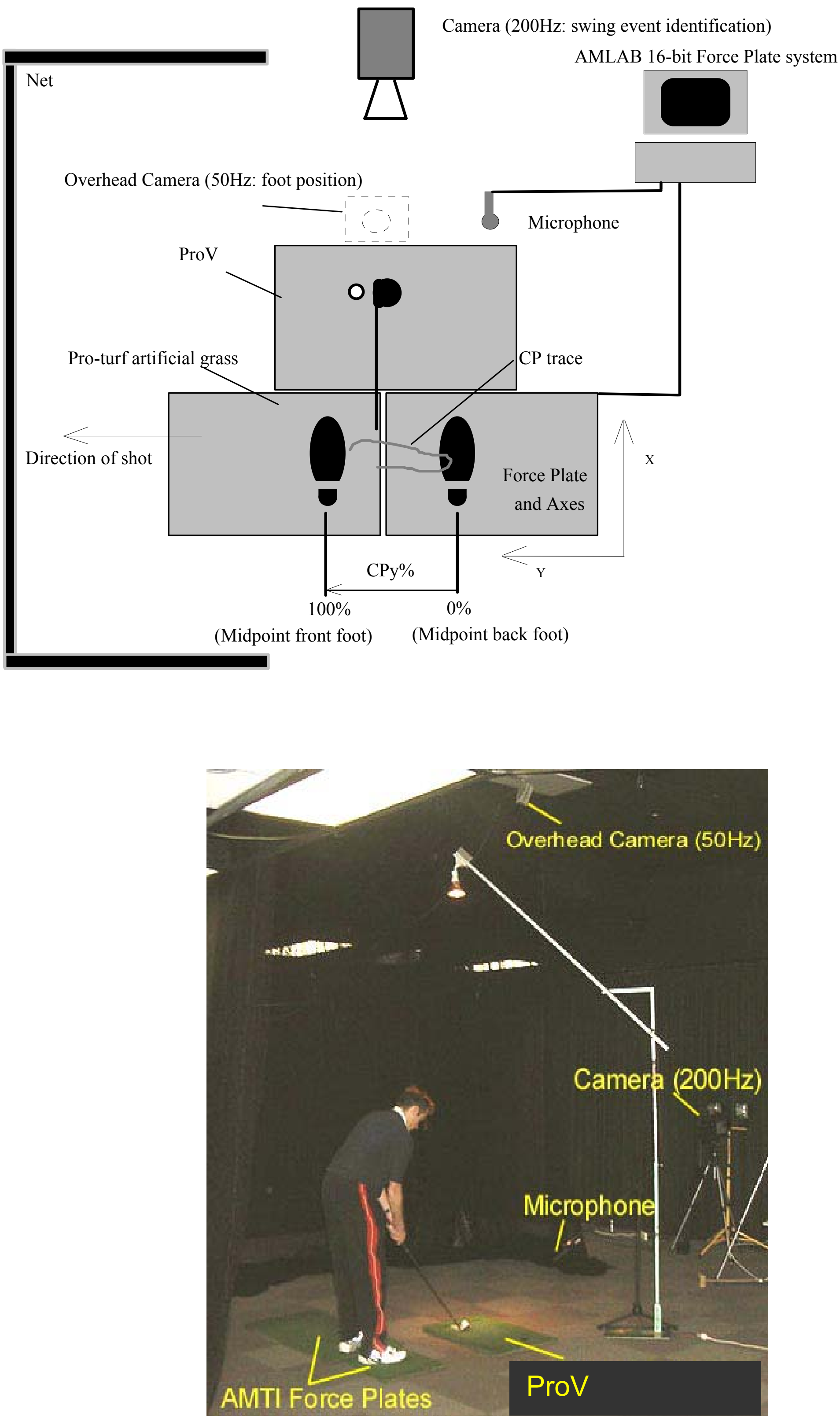


Figure 4.1: Laboratory set-up

4.2.4 Biomechanical analysis

4.2.4.1 Centre of pressure

Centre of pressure (CP) was chosen to represent weight transfer. Previous research has used CP or vertical forces (Fz). Comparison of CP and Fz data parallel to the line of shot in this study indicated that the measures were very similar (r = 0.999, $p < 0.001$, N = 62; Appendix B). As such, either measure would be appropriate (refer to Literature review, page 4, and Appendix A for discussion of different measures).

4.2.4.1.1 MEASUREMENT OF FORCES TO CALCULATE CP

CP was calculated from data obtained from two AMTI force plates (AMTI LG6-4, 1200 mm x 600 mm; AMTI OR6-5-1, 450 mm x 550 mm) one under each foot. The force plates were positioned such that the golfer could adopt their preferred stance while each foot remained wholly on each force plate. Force and moment data collected during each swing was passed through AMTI amplifiers (AMTI SGA6-4 attached to the LG6-4 force plate and an AMTI SGA6-3 amplifier attached to the OR6-5-1 force plate) set at a maximum gain of 4000. The data was then passed through a 24.3 Hz low pass filter and sampled by an AMLAB 16-bit data acquisition system (AMLAB Technologies, Sydney, Australia) at 500 Hz.

A note on the sampling rate: 500 Hz was chosen as a balance between precision and size of data files for subsequent analysis procedures. As some of the measures used in this study were instantaneous (i.e. at a particular swing event or instant in time) the higher the sampling rate, the closer the measure could be made to the particular instant (i.e. improving precision). Obviously, the higher the rate of sampling, the smaller the potential error due to not sampling at the instant required. For example, with a 500 Hz sample rate, assuming the swing event was identified accurately, the sample closest to that instant in time would be within +/- 0.001 s, compared to +/- 0.005 s for 100 Hz sampling rate. Force plate sampling rates in previous research have ranged from 100 Hz (e.g. Williams and Cavanagh, 1983) to 1000 Hz (Barrentine *et al.*, 1994).

Data was sampled for 1 s prior to and 0.5 s after ball contact. A sensitive microphone, located unobtrusively near the hitting area, detected the sound of ball contact and controlled storage of force plate data. The microphone signal was passed through a Peak EBU system, amplified, and then passed to the AMLAB system, which used a preset threshold to detect the instant of ball contact.

4.2.4.1.2 CALCULATION OF CPy

CPy for each force plate was calculated parallel to the line of shot (CPy1 and CPy2; equations 4.1 and 4.2). Overall CPy was calculated using this data (equation 4.3) and then smoothed using a 15 Hz Butterworth digital low pass recursive filter (Winter, 1990; see below for a summary of the choice of smoothing cut-off. Complete discussion in Appendix C).

Force Plate 1 (OR6-5-1, back foot for right handed golfers)

$$CPy1 = \frac{Mx1 - (Fy1 * Dz1)}{Fz1} \qquad \text{Equation 4.1}$$

Force Plate 2 (LG6-4, front foot for right handed golfers)

$$CPy2 = \frac{Mx2 - (Fy2 * Dz2)}{Fz2} \qquad \text{Equation 4.2}$$

Overall CP (Force plate 1 and 2 combined)

$$CPy = \frac{(Fz1 * CPy1) + [Fz2 * (CPy2 + Df2)]}{Fz1 + Fz2} \qquad \text{Equation 4.3}$$

Where Mx = moment about the x-axis

Fy = force in the y-axis (horizontal)

Fz = force in the z-axis (vertical)

Df2 = distance between the centre of force plate 1 and 2 (centre of force plate 1 = zero)

Dz = distance between transducer and grass surface. Calculated by adding the synthetic grass thickness (0.0318 m) to the distance from the force plate surface to the transducer (specified by the manufacturer)

For force plate 1 (AMTI OR6-5-1): Dz1 = 0.0353 m + 0.0318 m = 0.0671 m

For force plate 2 (AMTI LG6-4): Dz2 = 0.0535 m + 0.0318 m = 0.0853 m

The smoothed overall CP data was then used to calculate CP velocity using a 3-point central differences method (Nakamura, 1993, equation 4.4: 3-point, 5-point and 9-point difference methods all produced similar results so the simplest was chosen). This data were smoothed again using a 15 Hz Butterworth digital low pass recursive filter. The double filter method (i.e. both displacement and velocity data smoothed) was used as it has been found to produce better results compared with smoothing displacement data only (Giakas and Baltzopolous, 1997).

CP Velocity at sample n (3 point central differences method)

$$VelCPy_n = \frac{CPy_{(n+1)} - CPy_{(n-1)}}{2 * t} \qquad \text{Equation 4.4}$$

where n = sample at which velocity is calculated

t = time interval between samples; 0.002 s (500 Hz) for this study

It was decided to use a 15 Hz cut-off for both displacement and velocity data for this study. This cut-off was chosen based on three levels of decision making as recommended by Ball *et al.* (2001). First, 15 Hz was indicated as optimal by two of three automated algorithms (Challis, 1999; Winter, 1990; Yu *et al.*, 1999). The Challis (1999) and Winter (1990) methods produced similar cut-offs for CPy of 15.2 and 15.0 respectively. The Yu *et al.* (1999) method returned substantially larger values of approximately 25 Hz, due to the large sample rate in this study (500 Hz; the

Yu *et al.* method is strongly influenced by sample-rate). Second, the influence of different cut-offs on parameters of interest (CPy% between the feet and CPy velocity at swing events as well as maxima and minima) was inspected. Large changes in parameter values were evident when cut-offs below 15 Hz were used. This level was considered to represent over-smoothing. Third, visual inspection of raw and smoothed curves indicated the 15 Hz cut-off provided smooth displacement and velocity curves without attenuating what was considered real data, in particular near the maxima and minima.

Post-hoc analyses required the use of Fz% in cluster analysis and for direct comparison with other studies. This was calculated using equation 4.5.

$$Fz\% = \frac{(100 * Fz2)}{(Fz1 + Fz2)} \qquad \text{Equation 4.5}$$

Where Fz1 = Fz under the back foot

Fz2 = Fz under the front foot

4.2.4.1.3 NORMALISATION OF CPy

To normalise between subjects, CPy was expressed as a percentage of the distance between the back and front foot (measured at the point in the swing just before TA). The mid foot position of each foot (midway between the heel and toe) was calculated (figure 4.2). CPy displacement was expressed as a percentage (CPy%) of the distance between the back foot (0%) and front foot (100%).

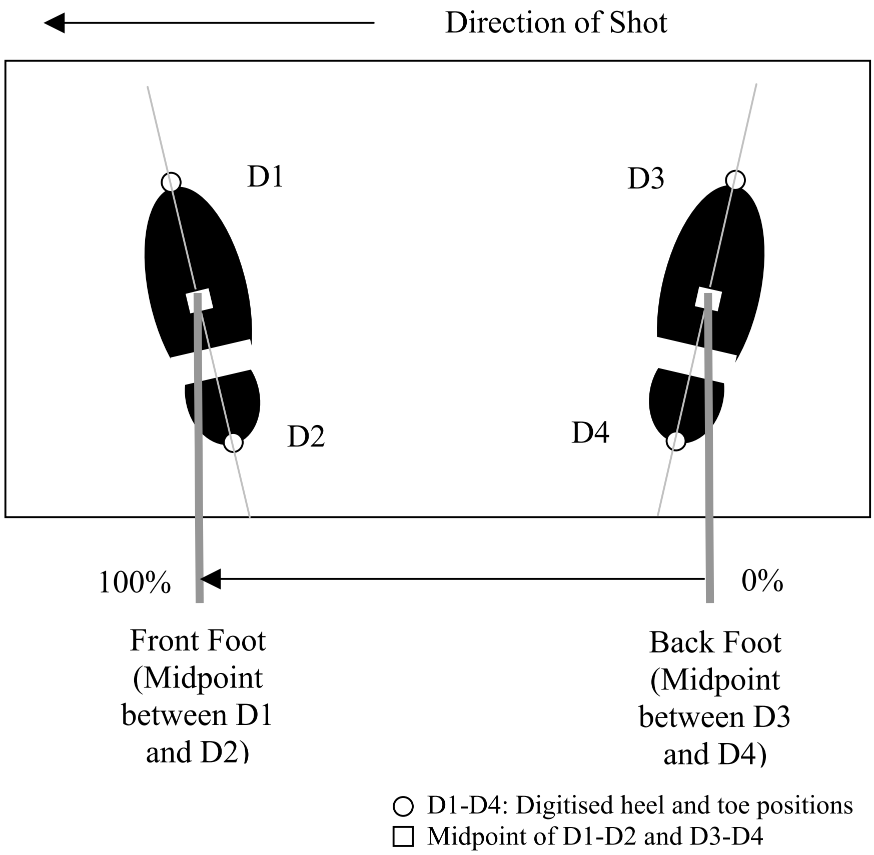
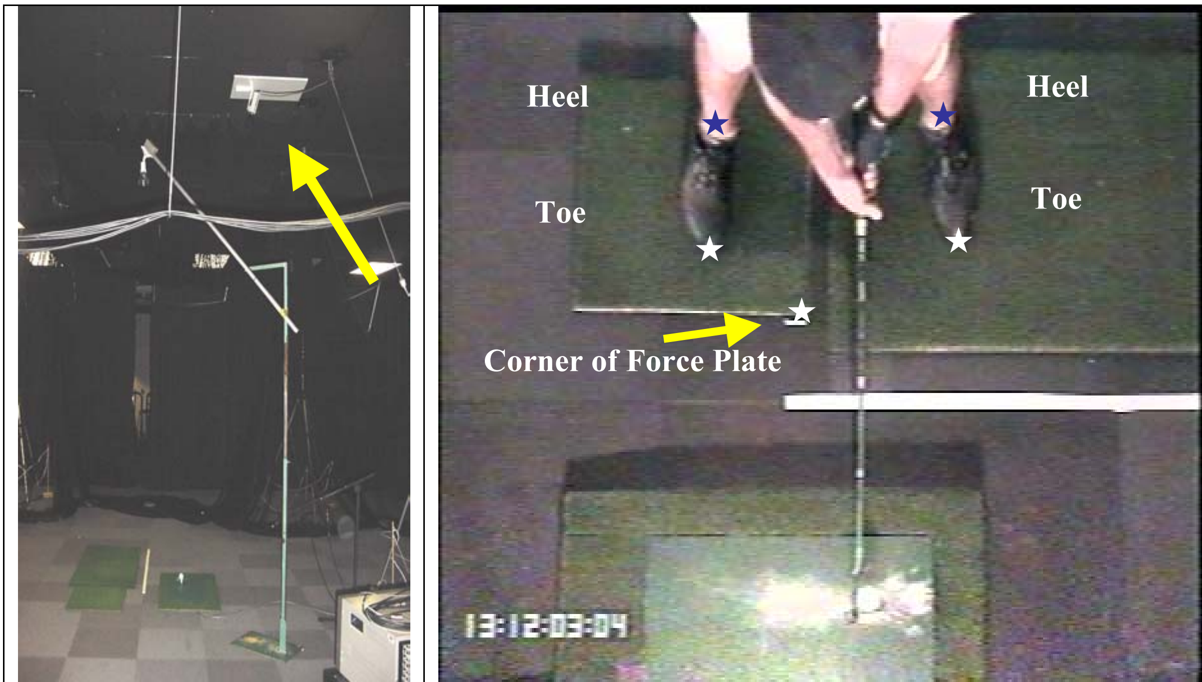


Figure 4.2: Normalisation of CPy (parallel to the line of shot).

4.2.4.1.3.1 Measurement of foot position

Foot position relative to the force plate (and hence relative to CP data), was obtained from video data. A 50 Hz Panasonic WV-CL350 camera was positioned above the hitting area [figure 4.3 (i)]. For each swing, the image just before TA was located and digitised using Peak MOTUS (Peak Performance Technologies Inc., Englewood, California). Four points were digitised to represent the feet; right heel, right toe, left heel, left toe. A fifth point, the corner of the small force plate, was digitised to allow for the combining of CP and digitised data [figure 4.3 (ii)].



(i) Position of overhead camera	(ii) View obtained from overhead camera (with digitised points denoted)

Figure 4.3: Foot position data collection. (i) Position of overhead camera relative to the hitting area and (ii) the view obtained from this camera (digitised points marked).

4.2.4.1.3.1.1 Perspective correction

The position of the overhead camera required offsetting to enable the feet to be seen past the upper body (refer to figure 4.3). To correct the out of plane image and the distortion produced in digitising, an adaptation of the algorithm presented by Begg *et al.* (1990) was used.

To calibrate and calculate the required perspective correction, a rectangular calibration board (0.96 x 0.72 m) was used. Prior to testing, the board was positioned over the hitting area, encompassing the position of the force plates and where the golfer's feet would be positioned. The axes of the calibration board were aligned with the force plate axes. The video image of the calibration board was recorded and the four corners of the calibration board (A, B, C and D; figure 4.4) were digitised from this image using Peak Motus. This was repeated four times, with the mean of the four trials used to represent the coordinates of the four corners.

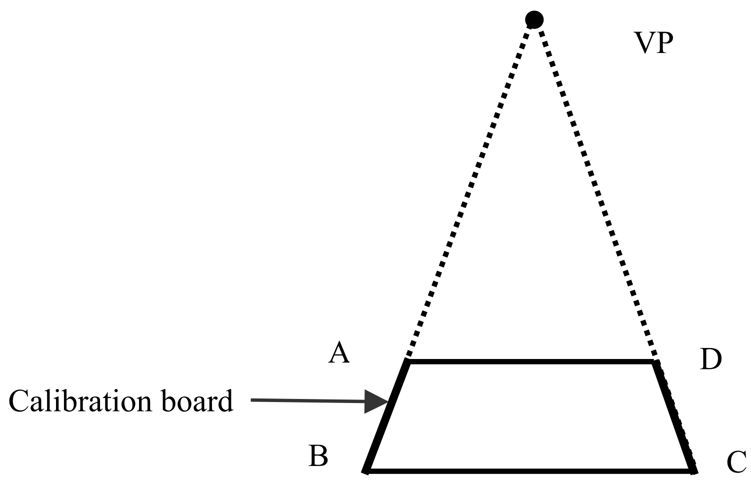


Figure 4.4: Location of the vanishing point (VP) using the sides of the calibration board (A, B and C, D).

Referring to figure 4.4, the two sides of the calibration board (line A-B and line C-D) were extended with the point of intersection defining the vanishing point (VP). The coordinates of A, B, C and D were shifted such that the middle of the calibration board was (0, 0). The vertical coordinate of VP was then established using equation 4.6.

$$\frac{u - u1}{u1 - u2} = \frac{v - v1}{v1 - v2}$$ Equation 4.6

where u = horizontal screen coordinate at VP

u1 = horizontal screen coordinate at A

u2 = horizontal screen coordinate at B

v = vertical screen coordinate at VP

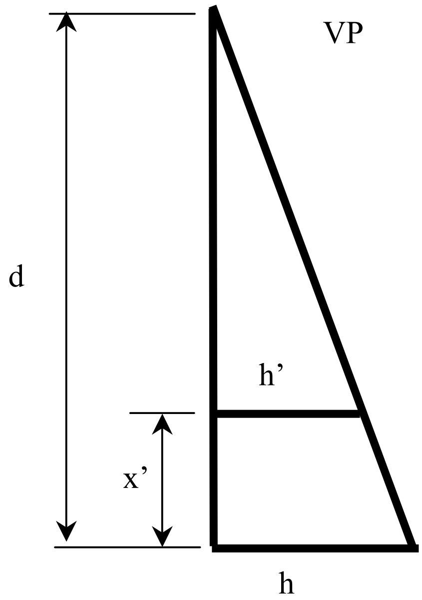
v1 = vertical screen coordinate at A

v2 = vertical screen coordinate at B

As the calibration device was in the centre of the screen with its vertical axis aligned with the vertical axis of the video, the horizontal coordinate of VP will be at the horizontal position on the screen; u = 0. Rearranging equation 4.6, the vertical coordinate of VP is obtained (equation 4.7).

$$v = \left(\frac{u - u1}{u1 - u2}\right) \times (v1 - v2) + v1$$ Equation 4.7

Referring to figure 4.5, using VP, perspective adjustment was made using the simple geometrical relationship (equation 4.8):



Perspective Equation*

h’ = h (1 – x’/d) **Equation 4.8**

$$h' = h\,(1 - x'/d) \qquad \text{Equation 4.8}$$

* Obtained from the relationship of the smaller and larger triangle i.e. the ratio of the large triangle d:h will be the same as the ratio of the small triangle (d-x’):h’

Figure 4.5: Geometrical relationship and equation used to adjust points due to perspective error.

4.2.4.1.3.1.1.1 Assessment of error in perspective correction

To assess the horizontal axis error of the perspective correction and calibration system for the foot digitising system, a board with a grid of 26 points was positioned over the hitting area and videoed using the overhead camera. These 26 points (with known coordinates) were digitised using Peak MOTUS. The digitised and known coordinates were then compared to indicate the error in the system (table 4.5).

Table 4.5: Comparison of known and digitised measures: horizontal screen axis (all measures in mm).

	Known Measure	Digitised Measure	Difference
1	0	0.0	0.0
2	239	238.7	0.3
3	479	479.7	0.7
4	720.5	720.9	0.4
5	117	115.8	1.2
6	355	352.9	2.1
7	598.5	596.7	1.8
8	0	-0.4	0.4
9	238	236.2	1.8
10	478	479.5	1.5
11	720	718.4	1.6
12	118	118.4	0.4
13	599	597.8	1.2
14	0	1.7	1.7
15	236	238.2	2.2
16	720	720.3	0.3
17	117	119.9	2.9
18	597	599.5	2.5
19	0	1.7	1.7
20	238	238.1	0.1
21	478	478.9	0.9
22	720	721.8	1.8
23	118	117.9	0.1
24	358	360.5	2.5
25	600	601.1	1.1
Mean			**1.25**
R.M.S.			**1.50**

Mean differences between measured and true coordinates were 1.25 mm with the maximum difference of 2.9 mm. Root mean square (R.M.S.) was 1.50 mm.

Note: These error values were not used to indicate overall error in the foot digitising process as it was considered more appropriate to indicate error based on actual testing processes. This process is outlined in section 4.2.5.3.

4.2.4.2 Measurement of swing events

A video operating at 200 Hz (Peak High Speed Camera; Peak Performance Technologies Inc., Englewood, California) with a shutter speed of 1/1000 s was placed facing the golfer and perpendicular to the line of shot (refer to figure 4.1). Each swing was recorded and a time code was overlaid on the image. The video was then used to identify eight swing events (table 4.6 and figure 4.6). This was performed by visual inspection with the aid of a grid placed over the monitor. The time at which each swing event occurred was recorded and then standardised to ball contact (i.e. ball contact = 0.0 s) to align with force plate data.

Table 4.6: Events in the golf swing used in this study.

	Event	Definition	Label
1	**Takeaway**	First backward movement of the club	**TA**
2	**Mid Backswing**	Club shaft parallel to the horizontal plane	**MB**
3	**Late Backswing**	Club shaft perpendicular to the horizontal plane when club is projected onto the YZ vertical plane	**LB**
4	**Top of Backswing**	Instant before shaft begins downswing	**TB**
5	**Early Downswing**	Club shaft perpendicular to the horizontal plane when club is projected onto the YZ vertical plane	**ED**
6	**Mid Downswing**	Club shaft parallel to the horizontal plane	**MD**
7	**Ball Contact**	Instant of club contact with ball	**BC**
8	**Mid Follow-through**	Club shaft parallel to the horizontal plane	**MF**

Note: the frame where the club was nearest perpendicular or parallel to the horizontal plane was used for MB, LB, ED, MD and MF.

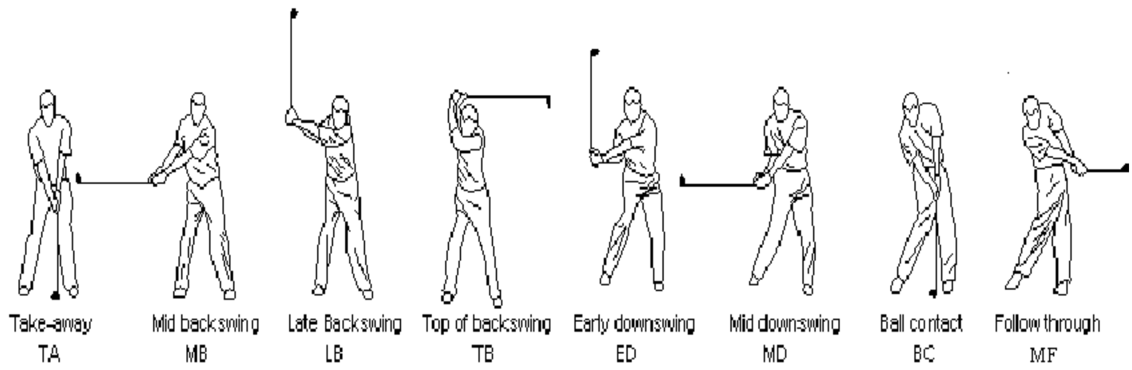


Figure 4.6: Golf swing events used in this study

The swing events TA, TB and BC were included as they represent key coaching events in the golf swing and have been used in previous studies. The choice of MB, LB, ED and MD was based on the criteria to increase swing events used but still provide easily identified events. As such it was decided to use the club to define positions and the most obvious events exist when the club is vertical and horizontal as viewed in the YZ plane. Initially the swing events Address and End of Follow Through were also quantified but identification of these events was unreliable and so the data was eliminated from further analysis.

The combination of 500 Hz force plate sampling and 200 Hz video data collection was not ideal as every second video data point lay between force plate data points. The initial test set-up used 50 Hz cameras. However, analysis of the errors associated with the 50 Hz camera identifying swing events indicated high-speed video was necessary. The only available high-speed camera was 200 Hz and technical and software constraints did not allow for the sampling rate of the force plate data to be altered from 500Hz.

4.2.5 Performance and CP parameters quantified

4.2.5.1 Club Velocity

For each trial horizontal Club Velocity immediately before ball contact was measured using a ProV Swing Analyzer (Golftek Inc., Lewiston, Idaho). Briefly, this system has an overhead light, which is detected by two rows of sensors located immediately behind the golf ball and approximately 0.05m from the golf ball. As the club head moves towards ball contact, the first row of lights is broken (the light is blocked by the club head), starting a timer. The timer is stopped when the second row of lights immediately before the ball are broken (known distance divided by time = velocity). The accuracy of the system was evaluated by comparing it with digitized data and in all test trials the difference between the systems was within the factory error specifications of +/- 0.1 $m.s^{-1}$ (Appendix D).

4.2.5.2 CPy Parameters

CP parameters calculated from force plate data and used in cluster analysis are summarised in table 4.7. These were also used later to compare cluster groups using ANOVA.

Table 4.7: CPy% between the feet at eight swing events – used in cluster analysis.

CPy% between the feet	
At each swing event	Relative to the distance between the feet (%)
CPy%TA	
CPy%MB	
CPy%LB	
CPy%TB	
CPy%ED	
CPy%MD	
CPy%BC	
CPy%MF	

CP velocities at eight swing events and other CP parameters were also calculated from force plate data and later used to compare cluster groups (table 4.8). Other CPy parameters are detailed in this table also.

Table 4.8: CPy Velocity at eight swing events and other CPy parameters used to compare cluster groups.

CPy Velocity	
At each swing event **VelCPyTA** **VelCPyMB** **VelCPyLB** **VelCPyTB** **VelCPyED** **VelCPyMD** **VelCPyBC** **VelCPyMF**	Instantaneous velocity ($m.s^{-1}$)
Other CPy parameters	
VMaxCPy	Maximum CPy Velocity ($m.s^{-1}$) between TA and MF
tVMaxCPy	Time of VMaxCPy relative to ball contact (s)
MaxCPy	Maximum CPy% - furthest position towards front foot between TA and MF (%)
tMaxCPy	Time of MaxCPy relative to ball contact (s)
MinCPy	Minimum CPy% - furthest position towards back foot (%) between TA and MF
tMinCPy	Time of MinCPy - relative to ball contact (s)
CPyR	Absolute CPy Range (m). Maximum CPy – Minimum CPy
CPyR%	Relative CPy Range (%). MaxCPy% - MinCPy%

Evaluation of CPy% at swing events was chosen in preference to using information from CP – time curves. In the literature, both methods have been used (Swing Events: Wallace *et al.*, 1990; Mason *et al.*, 1995: Normalised time; Barrentine *et al.*, 1994; Neal, 1998) and both have advantages and limitations. Swing events were chosen for three reasons. First, swing events are easily understood by coaches and players. For example, while the top of backswing event is easily identified, the position of the body at 76% of the swing from TA to BC is not (76% was the mean normalized time that TB occurred for golfers in this study). Second, there is growing evidence to suggest that using time-normalised data (e.g. normalised time between TA and BC) can have significant flaws because of issues of temporal dependency (e.g. Forner-Cordero *et al.*, 2006). The problem arises from the inherent assumption that there is no variability in the timing of events between TA and BC and that no rescaling occurs during the percentage conversion. Due to substantially different speeds of club movement near takeaway, the same normalised time can represent very different stages. For example, TB occurred between 69% and 80% of the total swing time from TA to BC. Comparison of normalized data between golfers, then, will be comparing very different stages of the swing; i.e. there is variability in the timing of events and rescaling does occur, rendering time-normalisation flawed. Third, the use of time-based data would have produced problems with uneven weighting of the cluster analysis performed in this study due to very strong correlations between datapoints, particularly where CP moves slowly (and so a large number of points will exist near a certain point). For example, near TA, CP moves relatively slowly. Cluster analysis using time-based data would be influenced by this by producing clusters that were determined by differences around TA rather than the whole swing. The use of swing events eliminated this problem.

4.2.5.3 Summary of error assessment

Table 4.9 presents the error associated with each parameter used in this study for a single measure and across 10 trials for each golfer. It is summarised here to allow easy reference for the reader with calculations for each parameter presented in Appendix E. Briefly, evaluation of error in the parameters used in this study was difficult due to the need to combine data obtained from three different measurement systems (force plate data, digitized data and timing data) and the lack of a gold standard for comparison. A combination of experimental and theoretical methods was used to determine an approximate error for each parameter. Force plate data error was established by comparing known CP data (a grid of known dimensions at each intersection outlined on the force plate) with measured CP data (force applied with a javelin tip to the force plate at each grid intersection) as used by Sommer *et al.* (1997) to evaluate CP error in Kistler force plates. Foot digitising error was established by comparing known foot position (heel and toe positions measured while the golfer was standing on the force plates) with digitised coordinates of the same golfer in the same position. Timing error was approximated by comparing club positions obtained from digitised data (e.g. mid backswing was identified as when the club was nearest the horizontal plane) and relating the time at which this occurred to the visual identification from the video screen as used in this study.

Table 4.9: Error estimates for parameters used in this study (200 Hz camera). All values +/-.

Performance	Approximate Single Measure	Approximate Across 10 trials	Approximate across 15 golfers (error in group mean)
Club Velocity (m.s^{-1})	0.1	0.04	0.01
CPy% between the feet			
At downswing events			
CPy%TA	0.6	0.2	0.05
CPy%MB	0.7	0.2	0.05
CPy%LB	0.6	0.2	0.05
CPy%TB	0.7	0.2	0.05
CPy%ED	1.0	0.3	0.08
CPy%MD	0.7	0.2	0.05
CPy%BC	0.6	0.2	0.05
CPy%MF	1.0	0.3	0.08
Average	0.7	0.2	0.05
CPy Velocity (m.s^{-1})			
At downswing events			
VelCPyTA	0.18	0.05	0.01
VelCPyMB	0.18	0.04	0.01
VelCPyLB	0.18	0.04	0.01
VelCPyTB	0.18	0.06	0.02
VelCPyED	0.23	0.15	0.04
VelCPyMD	0.20	0.10	0.03
VelCPyBC	0.18	0.04	0.01
VelCPyMF	0.23	0.16	0.04
Average	0.19	0.08	0.02
Other CP parameters			
VMaxCPy (m.s^{-1})	0.18	0.06	0.02
tVMaxCPy (s)	0.001	< 0.001	
MaxCPy% (% between the feet)	0.6	0.2	0.05
tMaxCPy% (s)	0.001	< 0.001	
MinCPy% (% between the feet)	0.6	0.2	0.05
tMinCPy% (s)	0.001	< 0.001	
CPyR (m)	0.007	0.002	0.001
CPyR% (% between the feet)	0.8	0.3	0.08

Average error across ten trials for CPy% data was 0.2 % and for CPy velocity was 0.08 m.s^{-1}. Similar values were indicated for maximum and minimum parameters. For group means, the approximate error will be further reduced. To provide an indication of the error in group means, the approximate error of the mean for fifteen golfers was calculated and was 0.05% (using quadrature summation – the square root of (0.2% x 15) divided by 15. See Appendix E for full calculations). However a more conservative approach to reporting data was taken with CPy% reported in units of 1%

and velocity in units of 0.1 $m.s^{-1}$ with data reported to another decimal place if required.

4.2.6 Statistical analysis

For each individual, a mean value was obtained across the 10 swings for Club Velocity at ball contact and each CP parameter defined in tables 4.7 and 4.8. These mean values for each individual were then used in further analysis (e.g. group means, cluster analysis). As the analysis proceeded, a number of parameters not initially included were required (e.g. foot width data, swing time from TA to BC). The mean for each individual golfer across the 10 trials was the value used in each case.

4.2.6.1 Identification of swing styles: cluster analysis

To examine if different weight transfer styles existed, cluster analysis was performed using CPy% data at the eight swing events in SPSS version 10. The clustering process was performed hierarchically then repeated using the cluster means from this analysis as seeds in a non-hierarchical procedure (K-means cluster). Milligan (1996) reported this process to obtain clusters more reliably.

4.2.6.1.1 HIERARCHICAL CLUSTER ANALYSIS

Hierarchical cluster analysis was performed using the squared Euclidean distance dissimilarity measure and the between-groups linkage clustering strategy. As CPy% data was measured on the same scale, no standardization was required.

The agglomerative schedule and dendrogram were employed to firstly determine if clusters existed in the data and secondly, to decide on the possible optimal solution (number of clusters in the data). If there were large jumps in the agglomerative schedule and dendrogram, it can be considered that clusters may exist in the data and analysis continued (Gower, 1975). Conversely, if no jumps exist, then it is unlikely that there are clusters in the data. The first reasonably large jump in the agglomerative schedule and dendrogram was identified. All cluster solutions below this level were analysed further. For example, if the 10-cluster solution was the first jump in the data, then the 10-cluster to 2-cluster solutions were analysed. Cluster means within each cluster solution were calculated and used as seeds (group or cluster means) in the next stage of the analysis (non-hierarchical analysis).

Note: What constitutes a 'large' jump in the agglomerative schedule and dendrogram is not well defined (e.g. Everitt, 1979) and is decided in an ad-hoc manner. For this study, a conservative approach was adopted to ensure the optimal solution was not overlooked. For example, if the six cluster solution was optimal but a smaller (but still reasonably large) jump existed at the eight cluster solution, all clusters from the eight cluster to two cluster solutions were reanalysed and statistical analysis to determine the optimal solution was applied to all solutions. This was distinct from only choosing cluster solutions producing the larger jumps.

4.2.6.1.2 NON-HIERARCHICAL CLUSTER ANALYSIS

Non-hierarchical analysis was performed using the k-means cluster method in SPSS 10. Each cluster solution below the cut-off identified in the hierarchical process was analysed. For each, the seeds, or group means, obtained from the hierarchical analysis provided the starting point for the analysis and each golfer was clustered with the nearest seed.

Note: SPSS 10 offers only the Euclidean distance measure for this analysis, as opposed to the squared Euclidean distance measure used in the hierarchical cluster analysis. This researcher was concerned that this may have produced an inconsistency in the analysis. However, 100% of golfers clustered into the same groups when the analysis was repeated using the squared Euclidean distance measure in a custom developed Microsoft Excel spreadsheet. As such, the different distance measures did not affect the analysis.

4.2.6.1.2.1 Number of clusters

As there is no widely accepted method for deciding on the number of clusters in an analysis (e.g. Hair *et al.*, 1995), a number of techniques were used to substantiate the analysis as recommended by Milligan and Cooper (1985).

4.2.6.1.2.1.1 Statistical methods

Milligan (1996) recommended the use of two or more statistical methods for choosing the number of clusters in a dataset. As such, all non-hierarchical solutions were analysed using two stopping rules. Both compare the distances between cases within a cluster to distances between cases in different clusters but use different key parameters.

1. *Point Biserial Correlation*

   A larger correlation coefficient indicates a stronger relationship between cases within clusters compared with cases in different clusters. The optimal cluster solution was the one that returned the highest Point Biserial Correlation coefficient. It is calculated using the following formula (equation 4.9).

Point Biserial Correlation =

$$[\text{Mean (outside) - Mean (within)}] * \frac{\sqrt{\dfrac{proportion(outside) * proportion(within)}{Total(Outside + within)}}}{Overall.SD}$$

Equation 4.9

Where	Mean	=	mean distance between cases within each cluster (within) divided by the total number of distances or mean distance between each cluster (outside) divided by the total number of distances
	Proportion	=	number of distances between cases within each cluster divided by the total number of distances (within) or between cases in different clusters divided by the total number of distances (outside).

Overall SD = standard deviation of all distances between all cases

2. C Index

The lowest C-Index value indicated the optimal solution (equation 4.10).

$$\text{C-Index} = \frac{\text{Sum (D - all clusters) - Minimum (D)}}{\text{Maximum (D) - Minimum (D)}} \qquad \text{Equation 4.10}$$

Where D = distance between two cases.

These methods were chosen as Milligan and Cooper (1985) found them to be among the strongest methods for accurately determining the number of clusters in a data set. Formulas have been taken from the Milligan and Cooper paper.

The cluster solution was chosen if both methods indicated it was optimal. If there was no agreement between the stopping rules, the largest cluster solution (i.e. the one with the largest number of clusters) was chosen, as suggested by Milligan (1996).

4.2.6.1.2.2 Cluster validation

Similar to the decision on the number of clusters, there is no method that has been widely agreed upon for validation of clusters (e.g. Hair *et al.*, 1995). Once again, a number of methods were used to validate clusters in this study. These were:

1. *Point Biserial Correlation*.

   This was reported by Milligan (1981) to be one of the strongest methods of internal validation of cluster analysis. The use of this method in validation differs from its use as a stopping rule. As a stopping rule, the largest coefficient across all cluster solutions analysed indicated the optimal solution without regard for the strength of the relationship. For the validation of a cluster, the strength and significance level of the correlation are examined.

2. *Replication*.

   In this procedure, the cluster process was repeated with three randomly drawn subsets of N = 41, or two thirds of the data. This procedure examines the stability or robustness of clusters. The number of golfers who reclassify into the same clusters as they did in the original analysis is assessed, with a higher percentage of reclassification indicative of a more robust cluster (Hodge and Petlichkoff, 2000). A qualitative assessment of the similarity of the group mean patterns was also examined.

3. *Leave-one-out reclassification.*

This technique eliminates a golfer from the analysis, re-calculates the cluster group means and then re-clusters the golfer using the nearest neighbour method. Successful reclassification (i.e. the removed golfer is allocated to the same cluster) indicates robustness of the solution. An unstable cluster will be influenced by single golfers and will perform poorly in reclassification.

4. *One way ANOVA.*

Cluster groups were compared to identify significant differences between parameters used in clustering (internal) as well as parameters not used in clustering (external).

4.2.6.1.2.3 Theoretical considerations

The final level of decision-making in terms of number of clusters and validity of the cluster groups, and overall philosophy of the analysis, was based on finding the minimum number of meaningful clusters in the data (similar to the guiding philosophy in other studies such as Wilson and Howard, 1983; Forwood *et al.*, 1985).

4.2.6.1.3 CLUSTER SUMMARY

Figure 4.7 provides a graphical summary of the cluster process.

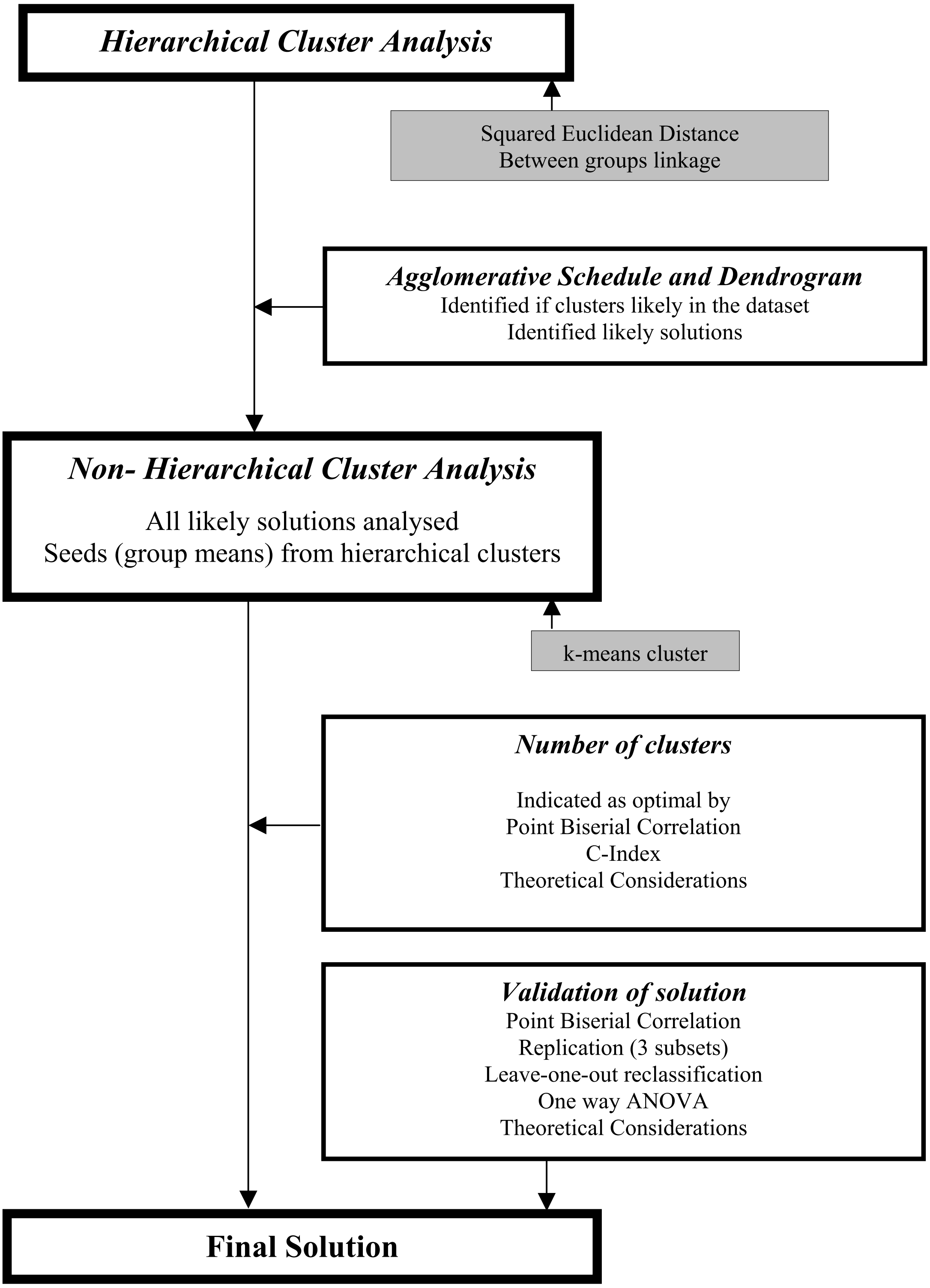


Figure 4.7: Summary of clustering process

4.2.6.2 Post-hoc statistical analysis - differences between clusters

To explore the differences between clusters, one-way ANOVAs were conducted on all remaining descriptive data (Age, Height, Mass and Handicap). Also, one way ANOVA of Club Velocity and CPy parameters (from table 4.7 and 4.8) used in the validation process of the cluster analysis were also used to compare groups descriptively. Effect sizes were calculated for all comparisons and used the subjective scale presented by Cohen (1988: $\eta^2 \geq 0.2$ – small, $\eta^2 \geq 0.6$ – medium, $\eta^2 \geq 0.12$ – large).

4.2.6.2.1 NON-NORMAL DATA

Pre-analysis screening using visual inspection of within-cluster histograms and Shapiro-Wilkes tests of Normality indicated that some CP parameters exhibited non-normality. Although ANOVA and regression (next section) are affected minimally by violations to the assumption of normality (e.g. Regression: Pedhazur, 1997; ANOVA: Tabachnick and Fidell, 1996) as group sizes were different, it was felt appropriate to assess the influence of the non-normal data.

The first method attempted to address the non-normal issue was to transform the data using various common transforms such as Log and inverse. However, this method was discounted as not all parameters could be successfully transformed. Further, the transforms made analyses difficult to interpret for coaching application.

The second method attempted, and the one used for this study, was the 'computer intensive' method described by Aron and Aron (1999). This method has been termed the 'within-dataset distribution' for this study as it generates a distribution from the dataset. After the original cluster analysis, group sizes for the clusters were defined. Maintaining these group sizes, golfers were randomly assigned to one of the groups to form a new 'random' grouping for ANOVA analysis. The ANOVA was repeated and the *F*-ratio recorded. This was repeated 1000 times (i.e. 1000 random datasets were constructed) then the *F*-ratios were sorted in ascending order to determine the 25th (lower 2.5%) and the 975th value (upper 2.5%).These values relate to a two tailed alpha level of $p = 0.05$. If the original *F*-ratio lay outside of these outer values (i.e. from 1-24 or from 976 to 1000), then this supports the relationship being a true rather than a random effect due to non-normal data. The upper or lower 0.5% ($p = 0.01$) value was also calculated.

4.3 RESULTS

4.3.1 Cluster analysis

4.3.1.1 Hierarchical cluster analysis: agglomerative schedule and dendrogram

Table 4.10 shows selected sections of the agglomerative schedule for the hierarchical cluster analysis of CPy% at eight swing events ($N = 62$ golfers). The cluster solution and jump in the coefficient columns have been added for easier interpretation.

Table 4.10: Selected sections of the agglomerative schedule for hierarchical cluster analysis of CPy% at eight swing events (N = 62 golfers).

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	62	144	-				
2	61	148	3	32	31	781	104
3	60	177	29	33	30	800	19
4	59	229	52	34	29	815	14
5	58	254	25	35	28	825	10
6	57	254	0	36	27	831	6
7	56	256	2	37	26	848	17
8	55	262	7	38	25	884	36
9	54	282	19	39	24	911	26
10	53	287	5	40	23	999	88
11	52	292	5	41	22	1025	26
12	51	296	4	42	21	1098	73
13	50	303	7	43	20	1105	8
14	49	311	8	44	19	1115	10
15	48	327	16	45	18	1133	18
16	47	412	85	46	17	1259	126
17	46	442	30	47	16	1291	31
18	45	455	12	48	15	1297	6
19	44	468	14	49	14	1393	96
20	43	491	22	50	13	1464	71
21	42	530	40	51	12	1476	12
22	41	531	0	52	11	1711	**235**
23	40	552	22	53	10	1780	69
24	39	557	5	54	9	2121	**341**
25	38	569	12	55	8	2277	156
26	37	583	14	56	7	2392	115
27	36	586	3	57	6	2870	**478**
28	35	605	19	58	5	3283	**413**
29	34	634	29	59	4	4119	**836**
30	33	667	33	60	3	4537	**418**
31	32	678	11	61	2	5024	**487**

The largest jumps in the agglomerative schedule existed in forming the 2-cluster to 6-cluster solutions. In order, these were the 4-cluster solution (agglomerative schedule jump in coefficients = 836), 2-cluster solution (487), 6-cluster solution (478), 3-cluster solution (418) and 5-cluster solution (413). The first reasonably large jump (235) existed at the 11-cluster solution.

Figure 4.8 represents the dendrogram from the hierarchical cluster analysis of CPy% at eight swing events (N = 62 golfers).

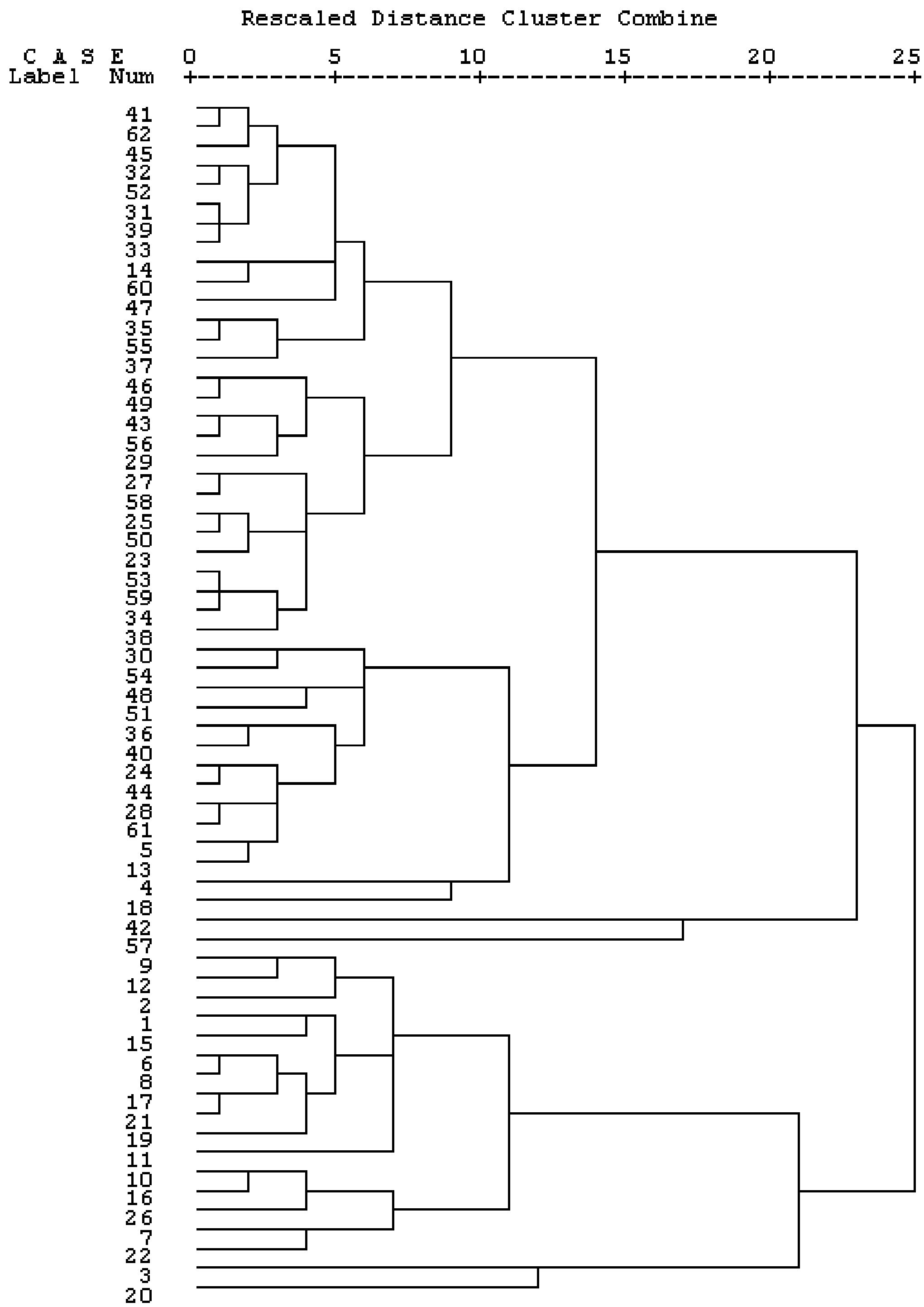


Figure 4.8: Dendrogram for hierarchical cluster analysis of CPy% at eight swing events (N = 62 golfers).

The larger jumps in cluster separation that existed in the agglomerative schedule were also evident graphically in the dendrogram. Examining the right hand side of the dendrogram, the horizontal distance between consecutive joins is relatively larger than those near the left hand side of the dendrogram. Also of note in the dendrogram were four individual cases (golfers) that clustered late in the hierarchical process. Cases 3 and 20 clustered together at the 6-cluster solution. Cases 42 and 57 clustered together at the 4-cluster solution.

It was decided to analyse all cluster solutions below the 11-cluster solution. This was based on the first reasonably large jump in the agglomerative schedule of 235, which corresponded with the 11-cluster solution, and the support from the dendrogram that no large jumps existed in higher solutions. As mentioned in the methods section 4.2.6.1.1, this was a conservative cut-off. While the likely solution would probably exist within the 2-cluster to 6-cluster solutions, evaluation of more cluster solutions at this point in the analysis was considered a safer approach towards finding the optimal cluster solution.

4.3.1.2 Hierarchical and non-hierarchical cluster solution means

Table 4.11 presents the CPy% means at the eight swing events for clusters within the 2-cluster solution to the 11-cluster solution for $N = 62$ golfers. These means were used as seed points in the non-hierarchical clustering process and means from this second analysis are presented in table 4.12.

Table 4.11: Mean CPy% at eight swing events for clusters within the 2-cluster to 11-cluster solutions from hierarchical analysis (N = 62 golfers).

Cluster	TA	MB	LB	TB	ED	MD	BC	MF
1 (N=18)	57	30	26	27	62	61	51	37
2 (N=44)	57	27	21	21	63	74	80	79
1 (N=18)	57	30	26	27	62	61	51	37
2 (N=42)	57	26	19	20	63	74	79	78
3 (N=2)	66	48	51	50	67	79	89	89
1 (N=16)	56	30	28	28	65	63	54	39
2 (N=2)	63	33	12	15	41	39	28	18
3 (N=42)	57	26	19	20	63	74	79	78
4 (N=2)	65	46	51	38	69	81	89	90
1 (N=16)	56	30	28	28	65	63	54	39
2 (N=2)	63	33	12	15	41	39	28	18
3 (N=42)	57	26	19	20	63	74	79	78
4 (N=1)	71	45	52	53	54	64	77	74
5 (N=1)	61	51	51	46	79	93	102	105
1 (N=16)	56	30	28	28	65	63	54	39
2 (N=2)	63	33	12	15	41	39	28	18
3 (N=14)	57	22	12	17	48	61	68	68
4 (N=28)	57	28	23	21	70	81	85	84
5 (N=1)	61	51	51	46	79	93	102	105
6 (N=1)	71	45	52	53	54	64	77	74
1 (N=16)	56	30	28	28	65	63	54	39
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=14)	57	22	12	17	48	61	68	68
4 (N=28)	57	28	23	21	70	81	85	84
5 (N=1)	76	15	8	15	43	31	25	22
6 (N=1)	61	51	51	46	79	93	102	105
7 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=14)	57	22	12	17	48	61	68	68
4 (N=5)	58	26	26	23	69	75	70	49
5 (N=28)	57	28	23	21	70	81	85	84
6 (N=1)	76	15	8	15	43	31	25	22
7 (N=1)	61	51	51	46	79	93	102	105
8 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=2)	51	28	9	8	40	52	53	49
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=5)	58	26	26	23	69	75	70	49
6 (N=28)	57	28	23	21	70	81	85	84
7 (N=1)	76	15	8	15	43	31	25	22
8 (N=1)	61	51	51	46	79	93	102	105
9 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=1)	51	16	8	8	51	59	49	38
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=5)	58	26	26	23	69	75	70	49
6 (N=28)	57	28	23	21	70	81	85	84
7 (N=1)	52	40	10	8	29	46	58	60
8 (N=1)	76	15	8	15	43	31	25	22
9 (N=1)	61	51	51	46	79	93	102	105
10 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=1)	51	16	8	8	51	59	49	38
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=5)	58	26	26	23	69	75	70	49
6 (N=14)	57	38	30	19	69	77	80	79
7 (N=1)	52	40	10	8	29	46	58	60
8 (N=1)	76	15	8	15	43	31	25	22
9 (N=14)	57	19	16	23	72	84	91	88
10 (N=1)	61	51	51	46	79	93	102	105
11 (N=1)	71	45	52	53	54	64	77	74

Table 4.12: Mean CPy% at eight swing events for clusters within the 2-cluster to 11-cluster solutions from non-hierarchical analysis (N = 62 golfers).

Cluster	TA	MB	LB	TB	ED	MD	BC	MF
1 (N=21)	56	30	23	25	59	59	51	38
2 (N=41)	58	27	22	22	65	76	82	81
1 (N=21)	56	30	23	25	59	59	51	38
2 (N=39)	57	25	19	21	64	76	81	80
3 (N=2)	65	46	51	38	69	81	89	90
1 (N=19)	56	30	25	26	61	62	53	41
2 (N=2)	63	33	12	15	41	39	28	18
3 (N=39)	57	25	19	21	64	76	81	80
4 (N=2)	65	46	51	38	69	81	89	90
1 (N=19)	56	30	25	26	61	62	53	41
2 (N=2)	63	33	12	15	41	39	28	18
3 (N=39)	57	25	19	21	64	76	81	80
4 (N=1)	71	45	52	53	54	64	77	74
5 (N=1)	62	47	51	31	77	89	96	98
1 (N=16)	55	30	27	27	65	62	52	37
2 (N=13)	57	22	11	16	50	62	68	68
3 (N=28)	57	28	23	22	69	81	86	85
4 (N=2)	63	33	12	15	41	39	28	18
5 (N=2)	63	42	45	44	56	70	77	67
6 (N=1)	61	51	51	46	79	93	102	105
1 (N=16)	55	30	27	27	65	62	52	37
2 (N=13)	57	22	11	16	50	62	68	68
2 (N=28)	57	28	23	22	69	81	86	85
4 (N=1)	51	50	16	16	40	47	32	13
5 (N=1)	76	15	8	15	43	31	25	22
6 (N=2)	63	42	45	44	56	70	77	67
7 (N=1)	61	51	51	46	79	93	102	105
1 (N=12)	54	31	27	29	62	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=12)	58	22	11	17	48	61	68	69
4 (N=8)	57	32	26	21	70	76	71	55
5 (N=26)	57	27	23	22	69	81	87	86
6 (N=1)	76	15	8	15	43	31	25	22
7 (N=1)	61	51	51	46	79	93	102	105
8 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=2)	51	28	9	8	40	52	53	49
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=10)	57	33	27	23	70	74	71	59
6 (N=23)	58	26	22	21	70	82	88	87
7 (N=1)	76	15	8	15	43	31	25	22
8 (N=1)	61	51	51	46	79	93	102	105
9 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=1)	51	16	8	8	51	59	49	38
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=10)	57	33	27	23	70	74	71	59
6 (N=23)	58	26	22	21	70	82	88	87
7 (N=1)	52	40	10	8	29	46	58	60
8 (N=1)	76	15	8	15	43	31	25	22
9 (N=1)	61	51	51	46	79	93	102	105
10 (N=1)	71	45	52	53	54	64	77	74
1 (N=11)	55	32	29	31	63	58	46	35
2 (N=1)	51	50	16	16	40	47	32	13
3 (N=1)	51	16	8	8	51	59	49	38
4 (N=12)	58	21	12	19	49	63	70	71
5 (N=6)	57	29	26	22	70	75	70	50
6 (N=13)	57	37	30	20	68	78	81	81
7 (N=1)	52	40	10	8	29	46	58	60
8 (N=1)	76	15	8	15	43	31	25	22
9 (N=14)	57	19	16	23	72	84	91	88
10 (N=1)	61	51	51	46	79	93	102	105
11 (N=1)	71	45	52	53	54	64	77	74

Of note in this data is the wide range of values evident in the BC and MF events, compared with other events. While other events showed differences of 20% to 40%, BC and MF ranged across 80% to 90%. This is particularly evident in the higher clusters (i.e. 11-cluster solution). Also of note is one subject produced a value greater than 100% at BC and MF (this is possible if CPy is moved to the outer edge of the front foot as 100% was defined at mid-foot).

4.3.1.3 Number of clusters

Table 4.13 reports results of Point Biserial Correlation and C-Index analysis on each of the cluster solutions in table 4.12. Repeating for clarity, the largest value for Point Biserial Correlation and the smallest value for the C-Index correspond to the optimal cluster solution.

Table 4.13: Point Biserial Correlation and C-Index data for each solution (*N* = 62 golfers). Largest value for each test in bold.

	11	10	9	8	7	6	5	4	3	2
Point Biserial Correlation	0.47	0.52	0.52	0.54	0.58	0.58	0.621	**0.622**	0.61	0.58
C-Index	3.48	1.66	1.66	1.48	1.28	1.27	1.26	**1.25**	1.34	1.37

Note: The 4-cluster and 5-cluster Point Biserial Correlation has been reported to 3 decimal places to make largest value clear.

Point Biserial Correlation and C-Index values indicated that the 4-cluster solution was optimal. This was the solution chosen for further analysis.

4.3.1.4 4-Cluster solution

Table 4.14 and figure 4.9 present mean CPy% at eight swing events for each cluster in the 4-cluster solution. Each cluster was labelled according to the pattern displayed to assist interpretation (refer figure 4.9).

Table 4.14: Mean CPy% at eight swing events for each cluster in the 4-cluster solution (*N* = 62 golfers).

Cluster	Label		TA	MB	LB	TB	ED	MD	BC	MF
1 (*N*=19)	Reverse	Mean	56	30	25	26	61	62	53	41
		SD	5	8	9	13	13	10	12	13
2 (*N*=2)	Extreme Back	Mean	63	33	12	15	41	39	28	18
	Foot Reverse	SD	17	24	6	1	1	11	6	7
3 (*N*=39)	Front Foot	Mean	57	25	19	21	64	76	81	80
		SD	5	11	11	9	12	5	11	11
4 (*N*=2)	Midstance Backswing	Mean	65	46	51	38	69	81	89	90
	Front Foot	SD	8	1	1	21	21	24	17	23

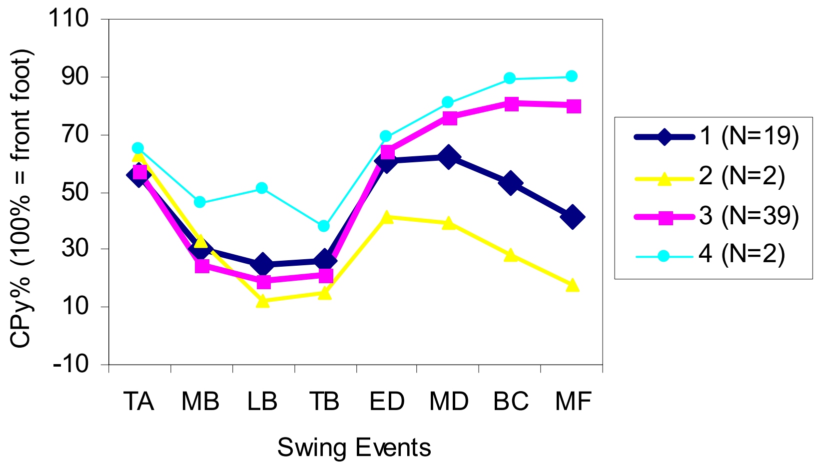


Figure 4.9: Mean CPy% for each cluster group at eight swing events (*N* = 62 golfers).

Cluster 1 and Cluster 3 formed the largest groups of *N* = 19 and *N* = 39 respectively. These clusters showed similar CPy% values at TA, MB, LB, TB and ED but Cluster 1 produced significantly smaller CPy% values at MD, BC and MF than cluster 3 ($p <$ 0.001; see section 4.3.2.2). Cluster 3 was labelled the Front Foot group as CPy% continued to increase towards the front foot from ED to MD and BC, with CPy% at BC (82%) positioned predominantly on the front foot. Cluster 1 was labelled the Reverse group due to the CPy% reducing (reversing in direction back towards the back foot) from MD to BC and continuing to reduce to MF.

Cluster 2 and Cluster 4 formed only small groups of $N = 2$. Cluster 2 was labelled the Extreme Back Foot Reverse group, as this group showed a similar pattern to the Reverse group, but CPy% values were smaller for most swing events, particularly at ED, MD, BC and MF, indicating CPy% was closer to the back foot. Cluster 4 was labelled the Midstance Backswing Front Foot group as CPy% was similar to the Front Foot group at ED, MD, BC and MF but was larger (i.e. less movement of CPy% towards the back foot in backswing or 'short backswing') at MB, LB and TB than the other groups and remained close to mid-stance. As can be noted in the dendrogram (figure 4.8), these four golfers did not combine to form clusters until late in the hierarchical clustering process. Cluster 2 formed at the 5-cluster solution (stage 58 of 62 in the hierarchical clustering process) and Cluster 4 formed at the 7-cluster solution (stage 56 of 62 in the hierarchical clustering process). Prior to these stages, the four golfers remained un-clustered with other golfers (i.e. formed groups of $N =$ 1).

Six golfers changed cluster groups between the hierarchical and non-hierarchical 4-cluster solution (Reverse to Front Foot – two golfers; Front Foot to Reverse – four golfers). The two smaller groups remained the same. The shift of some golfers into different groups was expected and is the reason for performing the non-hierarchical analysis to eliminate nesting in the hierarchical analysis. However, the low number of golfers who changed groups (<10%) indicated that the cluster solution was stable for the majority of the golfers.

4.3.1.5 Validation of clusters

4.3.1.5.1 POINT BISERIAL CORRELATION

The Point Biserial Correlation for the 4-cluster solution returned a large significant effect (r_{pbi} = 0.62, $p < 0.001$). Repeating for clarity, the strength (large effect) and significance level ($p < 0.001$) indicated the solution was valid.

4.3.1.5.2 REPLICATION

Repeating for clarity, cluster analysis on replication subsets was performed using the same procedures as reported in Methods section 4.4.2.6.1 and Results section 4.3.1.1 – 4.3.1.3. The results of replication analyses are summarized here. Agglomerative schedules and dendrograms are presented in Appendix F. Replication was repeated three times.

4.3.1.5.2.1 Replication subset 1

The first reasonably large jump in the agglomerative schedule corresponded with the 7-cluster solution. All cluster solutions below this cut-off were analysed non-hierarchically. Table 4.15 presents the Point Biserial Correlation and C-Index data for each non-hierarchical solution.

Table 4.15: Point Biserial Correlation and C-Index data for each solution for replication subset 1 (*N* = 41 golfers). Optimal value for each in bold.

	7	6	5	4	3	2
Point Biserial Correlation	0.56	0.56	0.59	**0.66**	0.65	0.63
C-Index	1.47	1.53	1.95	**1.41**	1.48	1.74

As Point Biserial Correlation and C-Index indicated that the 4-cluster solution was optimal, this solution was chosen. Table 4.16 and figure 4.10 present mean CPy% at each swing event for each cluster in the 4-cluster solution.

Table 4.16: Mean CPy% at eight swing events for each cluster in the 4-cluster solution for replication subset 1 (*N* = 41 golfers).

Cluster	TA	MB	LB	TB	ED	MD	BC	MF
1 (*N*=15)	56	28	26	27	63	62	53	40
2 (*N*=24)	57	23	16	21	62	74	81	83
3 (*N*=1)	71	45	52	53	54	64	77	74
4 (*N*=1)	51	50	16	16	40	47	32	13

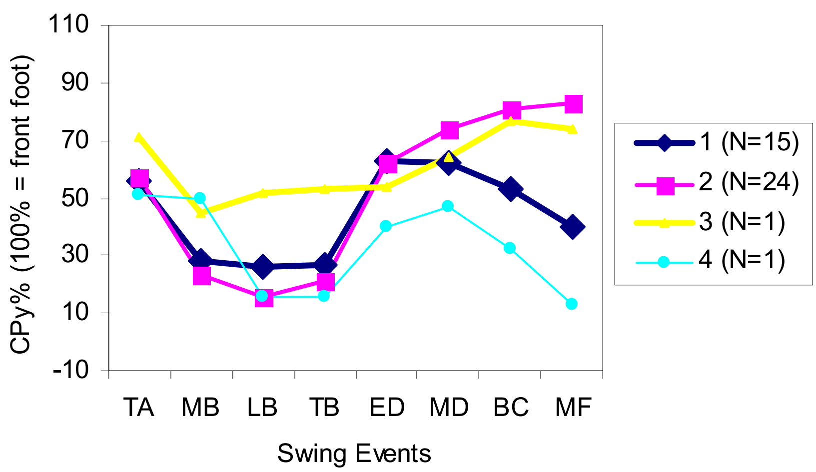


Figure 4.10: Mean CPy% at eight swing events for each cluster in the 4-cluster solution for replication subset 1 (*N* = 41 golfers). Note that the colours but not the group numbers correspond to the groups in the original analysis.

Cluster 1 and Cluster 2 were very similar in pattern to the Reverse and Front Foot groups respectively. Cluster 3 was formed by N = 1 golfers and was part of the Midstance Backswing Front Foot group while Cluster 4 was also formed by N = 1 golfers and was part of the Extreme Back Foot Reverse group. All golfers (100%) were reclassified into the same cluster as in the original analysis.

4.3.1.5.2.2 Replication subset 2

The first reasonably large jump in the agglomerative schedule corresponded with the 5-cluster solution. All cluster solutions below this cut-off were analysed non-hierarchically. Table 4.17 presents the Point Biserial Correlation and C-Index data for each non-hierarchical solution.

Table 4.17: Point Biserial Correlation and C-Index data for each solution for replication subset 2 (*N* = 41 golfers). Optimal value for each in bold.

	6	5	4	3	2
Point Biserial Correlation	0.55	0.56	0.57	**0.59**	0.57
C-Index	1.75	1.93	1.80	**1.24**	1.44

Both Point Biserial Correlation and C-Index indicated that the 3-cluster solution was optimal for subset 2 and was chosen for further analysed. Table 4.18 and figure 4.11 present mean CPy% at each swing event for each cluster in the 2-cluster solution from subset 2.

Table 4.18: Mean CPy% at eight swing events for each cluster in the 3-cluster solution for replication subset 2 (N = 41 golfers).

Cluster	TA	MB	LB	TB	ED	MD	BC	MF
1 (N=13)	53	31	28	31	60	58	48	32
2 (N=27)	57	27	20	21	61	73	78	74
3 (N=1)	76	15	8	15	43	31	24	22

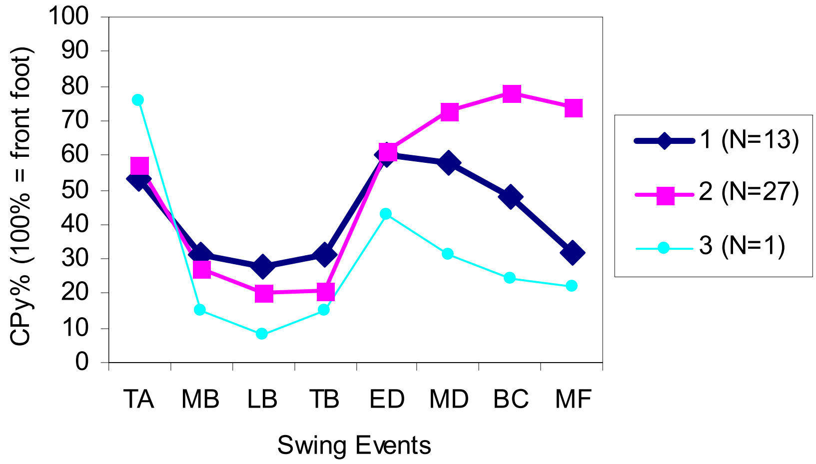


Figure 4.11: Mean CPy% at eight swing events for each cluster in the 3-cluster solution for replication subset 2 (N = 41 golfers). Note that the colours but not the group numbers correspond to the groups in the original analysis.

Once again, Cluster 1 and Cluster 2 were very similar in pattern to the Reverse and Front Foot groups while Cluster 3 was similar to the Extreme Back Foot Reverse group. All golfers (100%) reclassified into the same cluster as in the original analysis.

4.3.1.5.2.3 Replication subset 3

The first reasonably large jump in the agglomerative schedule corresponded with the 5-cluster solution. All cluster solutions below this cut-off were analysed non-hierarchically. Table 4.19 presents the Point Biserial Correlation and C-Index data for each non-hierarchical solution.

Table 4.19: Point Biserial Correlation and C-Index data for each solution for replication subset 3 (*N* = 41 golfers). Optimal value for each in bold.

	5	4	3	2
Point Biserial Correlation	0.55	0.55	0.59	**0.61**
C-Index	1.67	2.27	1.42	**1.19**

Both Point Biserial Correlation and C-Index indicated that the 2-cluster solution was optimal for subset 3. This solution was chosen for further analysed. Table 4.20 and figure 4.12 present mean CPy% at each swing event for each cluster in the 2-cluster solution.

Table 4.20: Mean CPy% at eight swing events for each cluster in the 2-cluster solution for replication subset 3 (*N* = 41 golfers).

Cluster	TA	MB	LB	TB	ED	MD	BC	MF
1 (*N*=13)	55	31	27	27	62	61	50	34
2 (*N*=28)	57	26	19	20	63	73	78	77

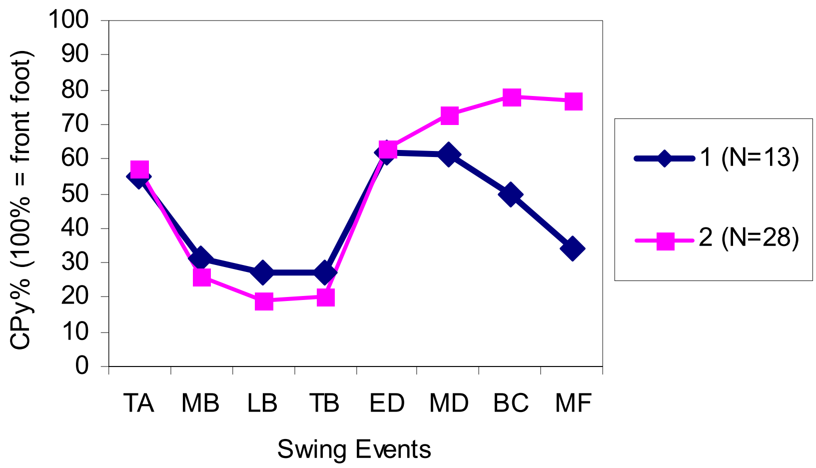


Figure 4.12: Mean CPy% at eight swing events for each cluster in the

2-cluster solution for replication subset 3 (*N* = 41 golfers). Note that the colours but not the group numbers correspond to the groups in the original analysis.

Cluster 1 and Cluster 2 were very similar in pattern to the Reverse and Front Foot groups respectively. All golfers reclassified into the same clusters as in the original analysis for the Front Foot and Reverse groups. However, two golfers from the small clusters did not reclassify into the same group. One golfer from the Midstance Backswing Front Foot group clustered with the Cluster 2 (Front Foot group) while one golfer from the Extreme Back Foot Reverse group reclassified with Cluster 1 (Reverse Group). The overall reclassification percentage was 95%.

4.3.1.5.3 Leave-one-out reclassification

Using the leave-one-out method of reclassification, 60 of the 62 golfers (97%) were re-classified into the same cluster group. Of the two golfers not to be re-classified correctly, one was from the Extreme Back Foot Reverse group, who was re-classified into the Reverse group and the other was from the Midstance Backswing Front Foot group who re-classified into the Front Foot group. The small groups showed instability in this classification.

4.3.1.5.4 One way ANOVA comparing groups

Due to small *N* in two clusters, only the two large clusters were compared. Table 4.21 reports Club Velocity and CP parameters for the Front Foot and Reverse groups as well as ANOVA results.

Table 4.21: Comparison of internal and external parameters between the Front Foot group and Reverse group.

	Front Foot group (N=39)		Reverse group (N=19)		ANOVA					Within-dataset distribution	
	Mean	SD	Mean	SD	F	p	Effect η^2	Effect Scale	Power	F (p=.05)	F (p=.01)
Internal Parameters											
CPy%TA	57	5	56	5	1.1	0.298	0.019	-	0.18	5.3	8.6
CPy%MB	25	11	30	8	2.0	0.163	0.027	Small	0.28	5.3	8.4
CPy%LB	19	11	25	9	2.4	0.132	0.041	Small	0.33	5.3	9.5
CPy%TB	21	9	26	13	3.1	0.081	0.046	Small	0.41	5.1	7.8
CPy%ED	64	12	61	13	0.8	0.368	0.009	-	0.14	5.2	7.4
CPy%MD	76	5	62	10	24.9	**<0.001**	0.312	Large	1.00	5.2	7.5
CPy%BC	81	11	53	12	90.5	**<0.001**	0.624	Large	1.00	4.8	8.6
CPy%MF	80	11	41	13	143.6	**<0.001**	0.721	Large	1.00	4.9	7.5
External parameters											
VelCPyTA	0.1	0.3	0.0	0.2	0.3	0.611	0.005	-	0.08	5.4	8.5
VelCPyMB	-0.3	0.2	-0.4	0.3	0.2	0.682	0.003	-	0.07	4.9	7.1
VelCPyLB	0.0	0.3	0.0	0.3	0.6	0.428	0.011	Small	0.12	5.7	9.3
VelCPyTB	0.3	0.5	0.1	0.4	2.2	0.148	0.037	Small	0.30	5.3	8.2
VelCPyED	1.1	0.7	0.6	0.9	**6.03**	**0.017**	0.097	Medium	0.67	5.5	7.8
VelCPyMD	0.9	0.8	-0.6	1.1	**35.3**	**<0.001**	0.387	Large	1.00	5.3	8.7
VelCPyBC	0.5	0.9	-1.2	1.1	**38.7**	**<0.001**	0.409	Large	1.00	5.1	7.5
VelCPyMF	-0.3	0.8	-0.3	1.1	0.01	0.906	0.000	-	0.05	5.8	10.1
VMaxCPy ($m.s^{-1}$)	2.5	0.7	2.5	1.0	0.0	0.924	0.001	-	0.05	4.6	7.0
tVMaxCPy (s)	-0.14	0.05	-0.18	0.14	2.8	0.101	0.054	Small	0.37	3.8	6.0
MaxCPy%	87	9	69	9	**52.3**	**<0.001**	0.480	Large	1.00	3.8	6.0
tMaxCPy% (s)	0.01	0.05	-0.23	0.24	**35.7**	**<0.001**	0.393	Large	1.00	4.7	6.1
MinCPy%	12	7	17	8	**5.7**	**0.020**	0.090	Medium	0.70	5.2	8.1
tMinCPy% (s)	-0.42	0.12	-0.41	0.21	0.1	0.792	0.003	-	0.06	5.3	8.2
CPyR (m)	0.36	0.07	0.27	0.08	**22.1**	**<0.001**	0.278	Large	0.99	4.4	8.7
CPyR%	75	11	51	13	**56.3**	**<0.001**	0.501	Large	1.00	5.2	7.2

Bold indicates significant at $p < 0.05$

Some internal (used in the clustering process, i.e. CPy% at eight swing events) and some external (not used in the clustering process) parameters were significantly different between the groups. The internal parameters CPy% at MD, BC and MF differed with large effects evident. External parameters CPy velocity at ED, MD and BC were different, with large effects evident for MD and BC. Also, external parameters CPy% maximum, time of CPy% maximum, CPy% minimum and CPy range in absolute and relative terms were also significantly different.

4.3.2 Comparison of Front Foot and Reverse cluster groups

This section refers to table 4.21 in the previous section.

Mean CPy% was significantly different between cluster groups for CPy%MD, CPy%BC and CPy%MF, with the Front Foot group producing the larger CPy% values. Effect sizes were large at all three events. The original *F*-ratios lay within the upper 0.5% of the within-dataset distribution *F*-ratios for all three comparisons indicating the significant differences were not influenced by non-normal distributions.

Mean CPy% was not significantly different between groups at all other swing events, with group differences less than 5%. However a small effect size existed between the groups for CPy%MB, CPy%LB and CPy%TB. Due to this small effect size, power was also low (<0.5).

Mean CPy velocity was significantly different between cluster groups for VelCPyED (difference = 0.53 $m.s^{-1}$), VelCPyMD (difference = 1.51 $m.s^{-1}$) and VelCPyBC (difference = 1.61 $m.s^{-1}$). The Front Foot group produced larger magnitudes of CPy velocity for VelCPyED and VelCPyMD (both positive) and smaller for VelCPyBC (negative). Effect size was medium for VelCPyED and the original *F*-ratio lay within the upper 2.5% of the within-dataset distribution *F*-ratios. Effect sizes were large for VelCPyMD and VelCPyBC and the original *F*-ratios lay within the upper 0.5% of the within-dataset distribution *F*-ratios for both comparisons indicating the significant differences were not generated by non-normal distributions.

Mean CPy velocity was not significantly different between groups at all other swing events. However a small effect size existed between the groups for VelCPyLB ($\eta^2 = 0.011$) and VelCPyTB ($\eta^2 = 0.037$). Due to this small effect size, power was also low in these comparisons.

VMaxCPy and tVMaxCPy were not significantly different between groups. However, a small effect size ($\eta^2 = 0.054$) existed for tVMaxCPy, with the difference in mean between the groups of 0.04 s. The Reverse group (-0.18 s) reached maximum CPy% velocity slightly earlier in the swing than the Front Foot group (-0.14 s). Power for this comparison was only 0.37.

Both MaxCPy% and tMaxCPy% were significantly different between the groups at $p < 0.05$ and returned large effect sizes. The Front Foot group produced a larger MaxCPy% and this occurred nearer to BC compared with the Reverse group. The original *F*-ratio for both parameters was in the top 0.5% of the within-dataset distribution *F*-ratios. However, tMaxCPy% produced groups with unequal variance (Levene's $p < 0.001$). This will be further analysed in section 4.3.2.4.1.

MinCPy% was significantly different between groups. The Reverse group exhibited a larger MinCPy% than the Front Foot group (difference = 5%), with a medium effect size ($\eta^2 = 0.090$). The original F-ratio for this comparison was in the top 2.5% of the within-dataset distribution F-ratios supporting the significant difference was not influenced by non-normal distributions. No difference was evident between groups for tMinCPy% with group means differing by only 0.01 s (practically no effect).

Both CPyR and CPyR% were significantly different between groups, with the Reverse group returning smaller values for both parameters (large effect for both differences between groups - CPyR = 0.10 m and CPyR% = 24%). The original F-ratio for both comparisons were in the top 1% of the within-dataset distribution supporting both significant differences were not generated by non-normal distributions.

Table 4.22 reports mean subject data, Handicap and Club Velocity for the Front Foot and Reverse groups, as well as the results of one-way ANOVAs comparing the groups. Only the upper limit of the within-dataset distribution F-ratio has been reported as the lower limit is not relevant to this analysis.

Table 4.22: Comparison of mean descriptive data, Handicap and Club Velocity for Front Foot and Reverse groups.

	Front Foot group (N=39)		Reverse group (N=19)		ANOVA					Within-dataset distribution	
	Mean	SD	Mean	SD	F	p	Effect η^2	Effect Scale	Power	F (p=.05)	F (p=.01)
Handicap	11.1	6.8	10.2	10.2	0.1	0.710	0.000	-	0.07	5.30	9.37
Age (years)	31.9	12.6	38.1	15.3	2.6	0.114	0.045	Small	0.36	5.72	9.24
Height (m)	1.80	0.06	1.81	0.07	0.1	0.827	0.001	-	0.05	3.28	3.86
Mass (kg)	80.4	8.6	83.3	10.0	1.3	0.269	0.022	Small	0.20	4.53	6.93
Club Velocity ($m.s^{-1}$)	44.1	3.9	44.1	4.9	0.0	0.981	0.000	-	0.05	1.89	2.56

Handicap showed no significant difference between the Front Foot and Reverse group. The difference in Handicap between groups (0.9) translated to an effect size of $\eta^2 = 0.003$, which was below the small effect ($\eta^2 = 0.02$) defined by Cohen (1968). The non-significant result was supported by the F -ratio falling well below the within-dataset distribution threshold of $F = 5.30$ ($p = 0.05$). As there was some ambiguity relating to whether Handicap should be considered ordinal or interval, a non-parametric test was also conducted to determine if the groups were significantly different. This test also returned a non-significant finding, supporting no difference between groups (Kruskall-Wallace: 0.59, $p = 0.441$). There was no significant difference between Height and Club Velocity with effect sizes for these variables also below $\eta^2 = 0.02$.

Age and Mass also showed no significant difference between groups at $p < 0.05$. However a small effect size existed for both variables (Age: $\eta^2 = 0.045$, Mass: $\eta^2 = 0.022$). The difference in mean Age between groups was 6.2 years while the difference in Mass was 2.9 kg. Power in both analyses was less than 0.5.

4.3.2.1 Post-hoc analysis of tMaxCPy%

Examination of the tMaxCPy% data indicated that the large variance in the Reverse group was due to one large and three moderate outliers. To further explore this relationship, a number of subsets of data were re-analysed until Levene's became non-significant. These were, in order:

1. Major outlier ($N = 1$) was removed from the Reverse group. All remaining data analysed (Reverse, $N = 18$; Front Foot, $N = 39$)
2. Group numbers evened using random deletion of cases, as recommended by Tabachnick and Fidell (1996; Reverse, $N = 18$; Front Foot, $N = 18$)
3. Major and minor outliers ($N = 4$) removed from Reverse group. All remaining data analysed (Reverse, $N = 15$; Front Foot, $N = 39$)
4. Group numbers evened using random deletion of cases (Reverse, $N = 15$; Front Foot, $N = 15$)

Table 4.23 reports the results of ANOVA using these subsets of data.

Table 4.23: Comparison of tMaxCPy% for Front Foot and Reverse groups with group sizes equal and outliers removed.

	ANOVA					
	F	p	Effect η^2	Effect Scale	Power	Levene's
1. N=1 outlier removed (Reverse N=18, Front Foot N=39)	45.6	<0.001	.455	Large	1.00	<0.001
2. Equal group numbers (N=18)	25.6	<0.001	.429	Large	0.99	<0.001
3. N=3 more outliers removed (Reverse N=15, Front Foot N=39)	44.5	<0.001	.460	Large	1.00	0.035
4. Equal group numbers (N=15)	24.5	<0.001	.468	Large	0.95	0.120

Based on the original F-ratio lying in the top 0.5% of within-dataset distribution F-ratios, a very strong p-value ($p < 0.001$) and a significant result in subset data where equal variance was achieved (table 4.23), the significant difference between the Front foot and Reverse groups for tMaxCPy% is supported.

4.4 DISCUSSION

4.4.1 Two weight transfer styles: description

Cluster analysis indicated that two major weight transfer styles existed. This result was supported by strong validation results. Point Biserial Correlation exhibited a large effect size and was significant at $p < 0.001$. Replication results were extremely strong, with both the Front Foot and Reverse groups appearing in all subset analyses. Further, 100% of large group cases reclassified into their original clusters (Front Foot or Reverse). Finally both internal and external parameters were significantly different between the groups, validating the two as separate styles.

Figure 4.13 shows the group CPy% means at eight swing events for the Front Foot group ($N = 39$), the Reverse group ($N = 19$) and All Golfers ($N = 62$).

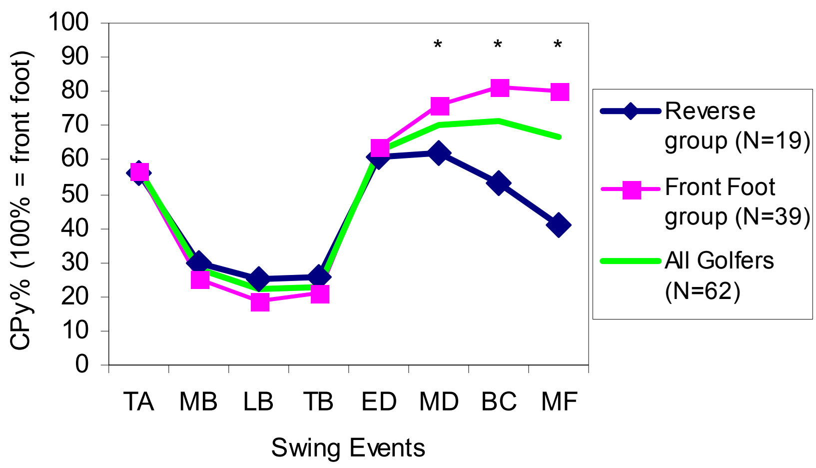


Figure 4.13: Group CPy% means at eight swing events for the Front Foot group ($N = 39$), the Reverse group ($N = 19$) and All Golfers ($N = 62$). Asterisks denote significant difference between Front Foot and Reverse groups.

The weight transfer pattern exhibited by the Front Foot group supported the coaching literature description of weight transfer. Referring to figure 4.13, CPy% was positioned approximately balanced between the feet at TA. During backswing, CPy% moved towards the back foot, ceasing this backwards movement just before the top of backswing, as indicated by a slightly lower CPy% value at LB (CPy%LB = 20%) compared with TB (CPy%TB = 21%). CPy% then moved towards the front foot between TB and ED, continuing towards the front foot to BC before moving slightly towards the back foot to MF. At BC, CPy% was positioned predominantly on the front foot (CPy%BC = 80%).

Weight transfer positions at swing events and overall pattern for the Reverse group did not support the coaching literature description of weight transfer. While CPy% was positioned midway between the feet at address, moved towards the back foot in backswing and then rapidly forward in early downswing, it did not continue onto the front foot at ball contact. From ED, the forward CPy% shift ceased as indicated by similar values for CPy%ED and CPy%MD (CPy% = 61% for both events). CPy% then moved towards the back foot from MD to a balanced position at BC (CPy% = 53%) and further towards the back foot at MF (CPy% = 41%).

4.4.1.1 Comparison of CPy% with the literature

Table 4.24 compares the reported CPy% positions for this study (All Golfers, Front Foot group, Reverse group) with other studies. Also, a meta-analysis including all available data for TA, MB, TB, MD and BC is presented. This was calculated by multiplying the reported parameter value by the N for each individual study, summing these multiplied values across all studies, then dividing by the total *N* across all studies. For example, for TA, two studies reported results (Koenig *et al.*, 1993; Robinson, 1994). Each value for TA was multiplied by N (For Koenig *et al.*, 1993: 14 x 55; For Robinson, 1994: Professional 10 x 58, Amateur 20 x 49). These values were summed and divided by the total N [i.e. (770+580+980)/(14+10+20)] to obtain a mean TA value for all studies. It should be noted that comparisons between the two styles found in this study and previous data are limited (and largely invalid) as previous studies did not account for different styles.

Table 4.24: Comparison of weight positions at different swing events in the golf swing for All Golfer (*N* = 62), the Front Foot group (*N* = 39) and the Reverse group (*N* = 19).

Study	Measure	Group	*N*	Swing Events							
				TA	MB	LB	TB	ED	MD	BC	MF
This Study	**CPy%**	**All Golfers**	**62**	**57**	**28**	**22**	**23**	**63**	**70**	**71**	**67**
This Study	**CPy%**	**Front Foot**	**39**	**57**	**26**	**20**	**21**	**64**	**76**	**82**	**81**
This Study	**CPy%**	**Reverse**	**19**	**56**	**30**	**25**	**26**	**61**	**61**	**53**	**41**
Cooper *et al.* (1974)	Fz%	Elite	5							75	50*
Richards *et al.* (1985)	COV%	Low HCP	10				28			96	
		High HCP	10				22			81	
Wallace *et al.* (1990)	CP%	Low HCP	1		53*		27		68*	82	
		High HCP	1		42*		31		47*	67	
Koenig *et al.* (1993)	Fz%	Low-High HCP	14	55			35				
Robinson (1994)	Fz%	Professional	10	58							
		Amateur	20	49							
Koslow (1994)**	'weight shift patterns'	Normal					27			62	
		Abbreviated					39			43	
		Reverse					60			36	
Meta Analysis (excluding Koslow)	All	All available		53 *N*=44	48 *N*=2		29 *N*=36		58 *N*=2	85 *N*=27	50 *N*=5

All values expressed as a percentage relative to the feet (0% = back foot, 100% = front foot). Transformed from the data presented in each study if required to allow for direct comparison between studies.

* MB, MD and MF have not been defined in these studies

** Koslow used set-up (assumed to be AD), top of swing (assumed to be TB) and BC and described the measure used only as 'weight shift patterns' with no further explanation.

TA and TB values for All Golfers, the Front Foot group and Reverse group in this study all lay within the range of values reported in other studies. TA values lay between the 58% (professional) and 49% (amateur) found by Robinson (1994). TB values for the three groups reported in table 4.24 in this study were larger than the 22% reported for High Handicap golfers by Richards *et al.* (1985) but smaller than all other results. Comparison between the meta-analysis indicated CPy% values at TA were slightly higher and values at TB were slightly lower than All Golfers in this study, although this difference was relatively small (4% for TA, 6% for TB).

CPy% values at BC for All Golfers, the Front Foot group and the Reverse group all lay within values previously reported in the literature (smallest = 36%, Koslow, 1994; largest = 96%, Richards *et al.,* 1985). However, a large discrepancy existed between the meta-analysis value at BC of 85% and the value for this study of for All Golfers of 71%. The lower value in this study was due to the influence of the Reverse group as the meta-analysis value and the Front Foot group value were similar (85% and 82% respectively).

Interestingly, the elite group tested by Cooper *et al.* (1974) reported a value that was only 4% different (75% compared with 71% in this study for All Golfers). The Cooper *et al.* study represents the only other study to report position of weight at BC for elite level golfers, with all other studies using either amateur golfers or not reporting values. These comparisons indicated that data in this study was reasonably similar to other studies. This comparison also highlights how the Reverse style could have existed but been undetected in other studies as values for All Golfers at TA, TB and BC all lay between the largest and smallest reported values in the literature.

The results reported by Koslow (1994) require further discussion. The ‘abbreviated’ group and the ‘reverse’ group reported by Koslow, along with the Reverse group in this study produced values at BC that were less than 60%. However, the patterns exhibited by the 'abbreviated' and 'reverse' groups in Koslow were dissimilar to the pattern exhibited by the Reverse group in this study. Koslow's 'abbreviated' group produced a larger value at TB compared with the Reverse group in this study (39% compared with 22%). Also, the difference between TB and BC was small for the 'abbreviated' group compared with the relatively larger difference for the Reverse group in this study (39% to 43% compared to 22% to 53%, difference = 4% compared to 31%). The 'reverse' group in Koslow differed substantially from the Reverse group in this study at TB and BC. Also, weight was nearer the back foot at BC compared to TB for Koslow's 'reverse' group compared with the Reverse group in this study, in which CPy% was nearer the front foot at BC. Also of note in this comparison was the golfers in the Koslow study were novices compared with this study which used subjects that at a minimum played five games of golf a year.

CPy% at MD was similar in this study compared with the values reported by Wallace *et al*. (1990) but CPy% at MB was smaller than MB in this study. However in both cases, the events were not defined in Wallace *et al.* and so the direct comparison might not be valid, as they might not refer to the same point in the swing.

No comparison data exists for the swing events LB or ED.

4.4.1.2 Comparison of CPy velocity with the literature

No comparison data exists for CPy velocity at swing events.

4.4.1.3 Comparison of other CPy parameters with the literature

There is little data in the literature for comparison with the other CPy parameters used in this study. This data is summarised in table 4.25.

Table 4.25: Comparison of other CPy parameters with the literature

Study	Measure	Group	*N*	Min	Max
This study	**CPy%**	**All Golfers**	**62**	**15**	**81**
This study	**CPy%**	**Front Foot**	**39**	**13**	**87**
This study	**CPy%**	**Reverse**	**19**	**18**	**69**
Richards *et al.* (1985)	COV	Low HCP	10	17	105
		High HCP	10	15	98
Koenig *et al.* (1993)	Fz%	Low-High HCP	14	20	
Meta-analysis	All	All available		18 *N*=34	102 *N*=20

MinCPy% for All Golfers in this study was the same as the high handicap group reported by Richards *et al.* (1985) but slightly smaller than the meta-analysis value. MaxCPy% was substantially smaller than Richards *et al.* and the meta-analysis although the meta-analysis included only the Richards *et al.* data. The 105% expressed by Richards *et al.* for low handicap golfers simply meant that COV% was positioned on the outside edge of the front foot (measurement of 100% was to the middle of the foot). However, it seems unusual that the mean value was this high as it suggested that all N = 10 golfers positioned weight near the outer edge of the base of support or that some produced values greater than 105%. In this study, only 6% of golfers (N = 4 of 62) produced a maximum over 100% with two golfers producing the maximum value recorded of 105% (one high handicap golfer and one social golfer). This difference between the studies might have been a function of different group samples or error in the Richards *et al.* measurement.

No comparison data exists for the other measures used in this study.

4.4.1.4 Reverse group weight transfer style in other studies

While the Front Foot style of weight transfer has been reported in coaching literature (e.g. Leadbetter, 1995) and scientific literature (e.g. Wallace *et al.*, 1990), the Reverse style has not. There are no qualitative reports of a similar style. Quantitatively, Cooper *et al.* (1974) reported a 'reverse' from BC (75%) to MF (50%). However, weight was positioned further towards the front foot at BC (75% compared with 53% for the Reverse group). Richards *et al.* (1985) reported large values at BC (96% and 81% for Low and High Handicap golfers respectively). Further, maximum weight position values occurred after BC. Neither finding supported a Reverse style of swing. Both golfers in the Wallace *et al.* (1990) study would have been in the Front Foot group in this study. None of the three weight transfer styles found by Koslow (1994) in novice golfers resembled the Reverse group.

While there is little indication of the Reverse group weight transfer style existing in previous qualitative and quantitative studies, this style may have been obscured in the averaging process. Using this study as an example, the predominant style in the data was the Front Foot group, with 63% of golfers exhibiting this style. As such, this would be the stronger influence in calculation of mean values and this would be reflected in the reported weight transfer path. Figure 4.14 shows the CPy% pattern using means from the whole group (*N* = 62) as well as the two major weight transfer styles identified in this study. This most resembles the Front Foot group with only a slight 'reverse' from BC to MF. As such, the Reverse style would have been obscured and the conclusion from the data would have been that the coaching defined weight transfer pattern was supported.

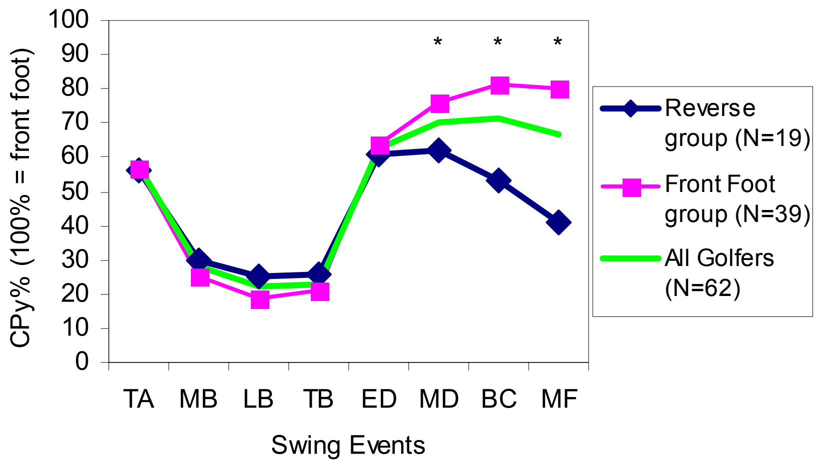


Figure 4.14: Comparison of mean CPy% at eight swing events between the Front Foot group (*N* = 39), Reverse group (*N* = 19) and All Golfers (*N* = 62). Asterisks denote significant difference between Front Foot and Reverse groups. Figure repeated here for clarity.

Another factor worth noting is the low number of swing events at which weight position was quantified in other studies. This has limited the assessment of weight transfer patterns and would be another reason why the Reverse style has not been reported previously. For example, if only TB and BC were used to quantify weight position, the Reverse group in this study would show movement from the back foot at backswing (CPy%TB = 26%) towards the front foot at BC (CPy%BC = 53%), although not as much as might have been expected from the previous literature.

Figure 4.15 shows the difference in patterns evident using all eight events compared with only TA, TB and BC. The use of more swing events would have made this weight transfer pattern more obvious if it did exist in previous studies. More importantly, even for the Reverse group, when analysing TA, TB and BC only, a Front Foot-like style would be concluded.

Figure 4.15: Comparison of the Reverse group pattern using all eight swing events and only TA, TB and BC.

4.4.2 Two weight transfer styles: comparison

4.4.2.1 Comparison of descriptive data between weight transfer styles

No difference existed between the Front Foot and Reverse groups for Handicap or Club Velocity. Further examination of these parameters indicated a fairly even spread of handicap levels and Club Velocities in both groups (figure 4.16 and figure 4.17). The Front Foot group included one professional and handicap ranges from 0 to 28. The Reverse Group included one professional and three amateurs involved in Australasian and US tour tournaments at the time of testing (and have since turned professional) as well as handicap ranges from 4 to 26. Interestingly, no social golfer was part of the Reverse group.

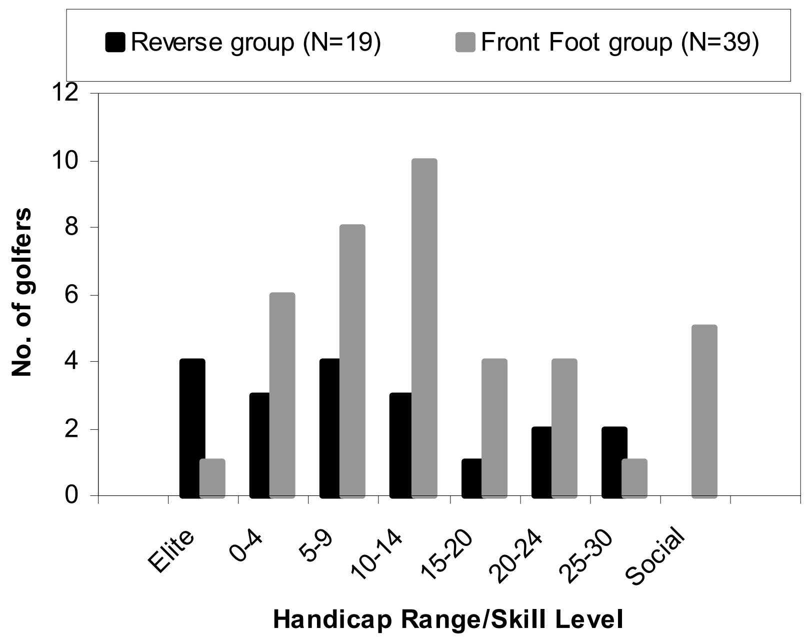


Figure 4.16: Distribution of Handicaps for the Front Foot and Reverse groups.

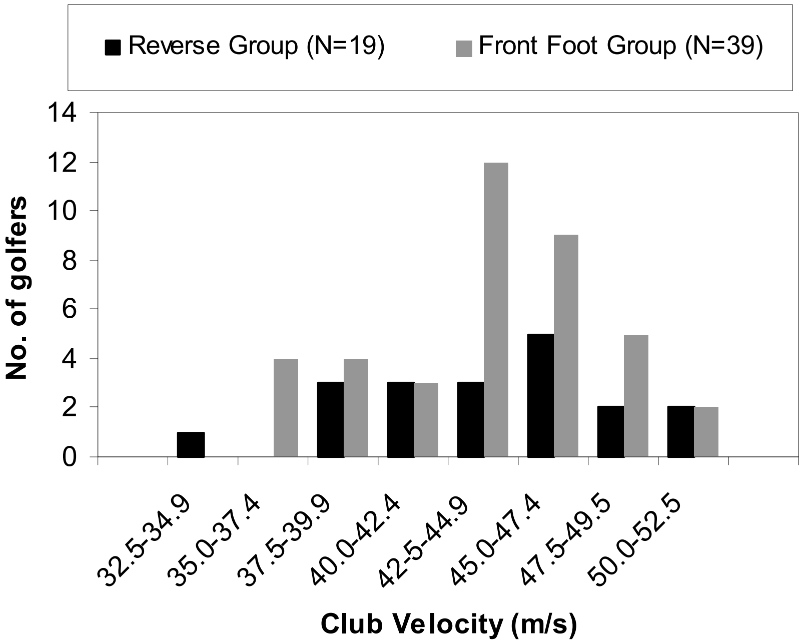


Figure 4.17: Distribution of Club Velocity for the Front Foot and Reverse groups.

The combination of statistical data and golfers included in each of the styles indicated that neither the Front Foot or Reverse styles were technical errors. There was no statistical difference in mean Club Velocity indicating skill level and performance were similar between groups and neither style was indicated as better than the other. Further, and importantly, both groups contained highly skilled golfers. The Reverse group (N = 19) comprised seven players (63%) who had a handicap less than five and included four players (21%) who were professionals or involved in Australasian and US tour tournaments at the time of testing. The Front Foot group contained one professional (3%) and seven golfers with handicaps less than five (18%). Establishing if the styles identified in this study are effective techniques is of particular importance in relation to the Reverse group, which has not been described in the coaching literature. Based on this combination of results, the effectiveness of both techniques is supported.

There was no difference between groups for Age, Mass or Height. This suggested that basic physical factors did not predispose a golfer to use a particular technique. It should be noted that there were small effects for Age (Reverse golfers were older – 38.1 years compared with 31.9 years) and for Mass (Reverse group = 83.3 kg, Front Foot group = 80.4 kg). However, as neither was significant, it would require more testing to determine if these factors are important in choosing one style or the other for a particular golfer. It would be useful in future work to identify the predisposing factors (if any exist) to using a particular style of swing. These might include appropriate strength and flexibility tests to more thoroughly assess if a physical attribute may influence the adoption of either the Front Foot or Reverse style of weight transfer.

4.4.2.2 Comparison of CPy% between weight transfer styles

Front Foot and Reverse groups were significantly different for CPy%MD, CPy%BC and CPy%MF. This indicated that downswing was a phase of major difference between the styles with the Front Foot group moving weight towards the front foot through BC, while the Reverse group moved towards the back foot after MD, balancing weight more evenly between the feet at BC (see figure 4.18, repeated for clarity). A small effect size also existed for CPy%MB, CPy%LB and CPy%TB. This indicated that the Reverse group tended to position weight in a more balanced position in backswing, although these differences were not significant. However, as will be discussed in section 4.4.2.4, minimum CPy, which occurred in backswing, was significantly different between groups.

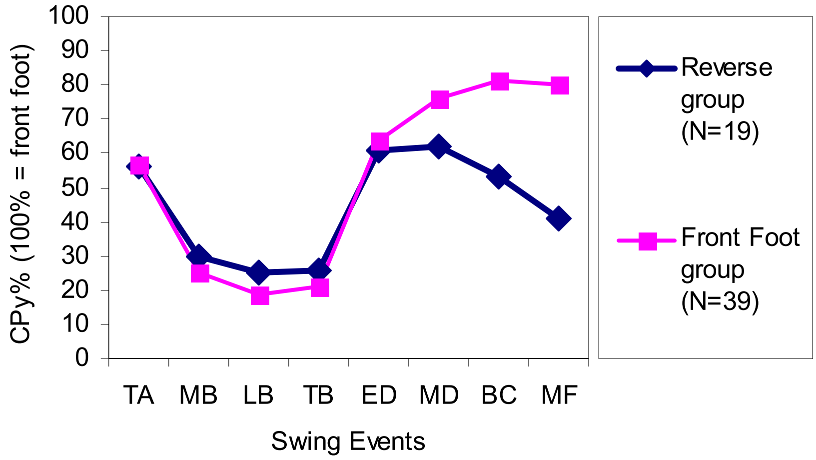


Figure 4.18: Mean CPy% at swing events for Front Foot and Reverse groups.

4.4.2.3 Comparison of CPy velocity between weight transfer styles

CPy velocity was significantly different between the Reverse and Front Foot groups at ED, MD and BC. As can be noted in figure 4.19, CPy velocity was smaller in magnitude for the Reverse group at ED while at MD and BC, it was moving in opposite directions. These differences were expected given the movement patterns of CPy% discussed in section 4.4.2.2. The results indicated that the Front Foot group was moving weight towards the front foot at TB and this movement direction continued until after BC. In contrast, while the Reverse group also moved weight towards the front foot at TB and ED, this group produced a significantly slower rate at ED compared to the Front Foot group, and moved weight towards the back foot at MD and BC. CPy velocities at TA, MB and LB were similar in magnitude indicating the major differences between the groups in terms of CPy velocity existed in downswing.

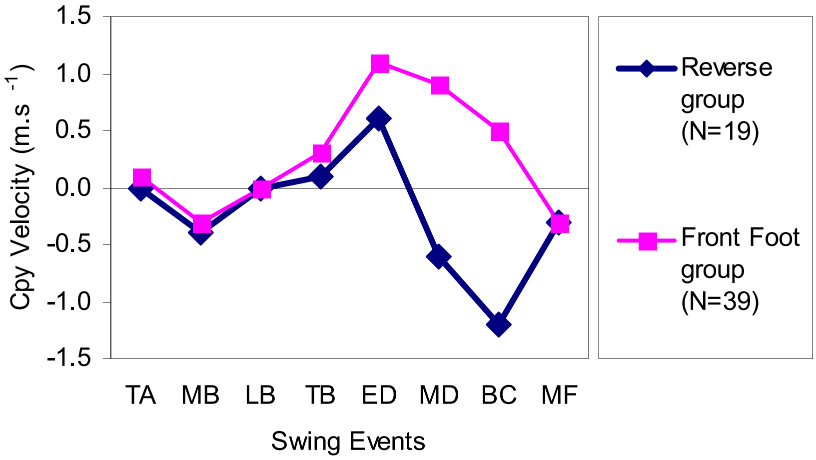


Figure 4.19: Mean CPy velocity at swing events for the Front Foot and Reverse groups.

Of interest are the similar mean values for VelCPyMF between the Front Foot and Reverse groups. While groups differed significantly at ED, MD and BC, both groups were moving CPy towards the back foot at MF at a similar rate. The value of 0.3 $m.s^{-1}$ for both the Front Foot and Reverse groups represents only approximately 12% of the absolute maximum velocities produced by either group so this movement is not rapid relative to other parts of the swing. A useful future direction for comparison of thee groups would be to identify a reliable swing event later in the follow through to examine if weight transfer patterns were once again similar later in the swing.

It should be noted that both groups produced similar maximum CPy velocity measures. This is mentioned as figure 4.19 can be misleading as it presents CPy velocity measures at swing events only. While it looks like the Front Foot group produced a greater positive CPy velocity compared to the Reverse group, the maximum CPy velocities produced by both groups were similar and occurred at similar times in the swing between TB and ED (see section 4.4.2.4 for discussion).

4.4.2.4 *Comparison of other CPy parameters between weight transfer styles*

The Front Foot group produced significantly larger CPyR and CPyR% values compared with the Reverse group. The Front Foot group moved CPy% through 75% of the distance between the feet (0.36m in absolute terms) compared with only 50% for the Reverse group (0.27m). Also, both MaxCPy% and MinCPy% were significantly different between groups, with the Front Foot group moving weight further towards the back foot in backswing (MinCPy% = 13% compared with 18%) and further towards the front foot in downswing (MaxCPy% = 87% compared with 69%) than the Reverse group. This indicated that the Front Foot group's larger range of weight transfer was produced at "both ends" of CPy movement (as opposed to a greater MaxCPy% position only, for example). Also it suggested that on average the Reverse group maintained the weight in a more central position between the feet throughout the swing, compared with the Front Foot group.

The time of maximum CPy% was significantly different between the Front Foot and Reverse groups with the Front Foot group achieving a maximum CPy% later in the swing. However, both groups achieved the minimum value at similar times. This further suggested differences between groups were most prominent in downswing. Maximum CPy velocity was not significantly different between the groups (2.5 $m.s^{-1}$ for both). Also, there was no difference in the time that this maximum occurred with a difference of only 0.04 s (Front Foot = -0.14 s, Reverse group = -0.18 s), although there was a small effect size between groups for this parameter. Examination of the data indicated that this maximum occurred between TB and ED for both groups. This indicated similarity in the nature of movement of weight in the early downswing phase for both groups.

4.4.2.5 Summary of differences between the Front Foot and Reverse styles

Both the Front Foot and Reverse groups positioned weight between the feet at takeaway, moved towards the back foot in backswing and started towards the front foot just before the top of backswing and into early downswing. From this point the Front Foot group continued to move towards the front foot, positioning weight predominantly on the front foot at ball contact. In comparison, the Reverse group after moving weight onto the front foot in early downswing, produced a 'reverse' weight shift such that weight was more balanced between the feet at ball contact. The rate of weight transfer between styles also differed in downswing with the Front Foot golfers producing large positive velocities towards the front foot in all downswing events compared with the Reverse golfers who produced negative velocities in mid downswing and ball contact. The Front Foot group produced a greater overall range of weight transfer, a larger maximum that occurred just after ball contact and a smaller minimum. This compared with the Reverse group which produced a smaller range of weight transfer, a smaller maximum that occurred in mid downswing and a larger minimum, indicating this technique maintained weight nearer a more balanced position through the swing.

4.4.2.6 *Post-hoc comparison between weight transfer styles in this study and Neal (1998)*

To further explore the two styles, post-hoc analysis was performed to compare this study with Neal (1998). Neal reported finding significant differences between the 'Right-to-Left' (possibly similar to the Front Foot group in this study) and 'Rotational' (possibly similar to the Reverse group in this study) for tMaxCPy and the ratio between CPxR and CPyR. To perform this comparison, tMaxCPy was recalculated as a percentage of the time between TA and BC (termed Time of Max CPy). CPx Range was calculated using the maximum and minimum CPx position between TA and BC (as the parameters were in metres, no normalization to foot position was required). Table 4.26 reports the results of the comparisons.

Table 4.26: Comparison of the Front Foot and Reverse groups with the Right-to-Left and Rotational groups from Neal (1998)

	Front Foot group (N=28)	Reverse group (N=12)	ANOVA			
	Mean ± s	Mean ± s	F	p	Effect η^2	Effect Scale
Time of Max CPy This study Neal study	 98 ± 4 (right to left: 99)	 87 ± 9 (rotational: 87)	 35.5	 <0.001 < 0.05	 0.40	 Large
Ratio CPxR:CPyR This study Neal study	 0.29 ± 0.12 (not reported)	 0.39 ± 0.14 (not reported)	 7.7	 0.010 < 0.05	 0.18	 Large

Both comparisons showed similarities. Time of Max CPy was significantly different between the groups in both Neal (1998) and this study. Also, similar Time of Max CPy values were returned for each group in both studies, with the Front Foot group returning similar values to the 'Right-to-Left' style golfers in the Neal (1998) study while the Reverse group was similar to the 'Rotational' style. Both studies also returned significant differences between groups for the ratio between CPx and CPy range. Both the Reverse group and 'Rotational' style produced significantly greater CPxR:CPyR ratio than the Front Foot group and 'Right-to-Left' style (Neal noted the 'Rotational' style produced a greater ratio but did not report values). Although more testing would be required to confirm this comparison, there is quantitative support for similarities between the two styles described in the Neal study and the two styles found in this study. This area shows excellent potential for further research.

4.4.3 General discussion of weight transfer styles

The existence of two different weight transfer pattern clusters does not support coaching literature on weight transfer during the swing which has largely considered only one style to exist (e.g. Leadbetter, 1995). Further, while the majority of the group tested (N = 39) exhibited a weight transfer pattern similar to that described in coaching texts of balanced at address moving to the back foot in backswing and to the front foot in downswing and through ball contact, not all golfers exhibited this pattern. Approximately one third of the golfers (N = 19), after moving weight forward from TB (CPy% = 26%) to ED (CPy% = 61%), then produced a reverse movement, such that the weight was positioned near midstance at BC (CPy% = 53%) and continued towards the back foot to MF (CPy% = 41%). There was no statistical difference in Handicap or Club Velocity between the groups and both contained highly skilled golfers. The Reverse group (N = 19) comprised seven players who had a handicap less than five and this seven included four players who were professionals or involved in Australasian and US tour tournaments at the time of testing. The Front

Foot group contained one professional and seven golfers with handicaps less than five. This suggested that neither pattern was a technical error.

Different styles within a skill or group being tested can reduce effect sizes, and hence power, in statistical analyses when all styles are assessed together. Briefly, type 1 or 2 errors can be produced by the existence of different styles or movement strategies in the skill being examined (Bates, 1996). For example, the correlation between CPyR and Club Velocity was not significant using the data as one group ($N = 62$, $r = 0.19$, $p = 0.134$) or for the Reverse group ($N = 19$, $r = 0.17$, $p = 0.185$) but was significant for the Front Foot group ($N = 39$, $r = 0.53$, $p = 0.001$). As such, if the data had been treated as one group, a type 1 error would have been made for the Front Foot group. This will be discussed in more detail in Study 2.

While two major clusters have been identified in this study, more work with larger subject numbers is required to examine if more clusters exist. As noted, one of the difficulties in the use of cluster analysis is the decision on how many clusters to use, as no widely accepted stopping rules exist (Hair *et al.*, 1995). This being the case, it would be presumptuous to suggest that only two styles exist. It may also be that golfers fit along a continuum of weight transfer styles, although the very strong validation results for the clusters would suggest that the styles found in this study do exist and are strongly defined. Further, using CPy%MF as an example (figure 4.20) distribution of all golfers showed two peaks and within the two peaks, the data was reasonably normally distributed for each cluster group. While this is a limited method of examination (cluster analysis is a multivariate technique and multidimensional graphics were not available to this researcher), it highlights two peaks in the distribution of all data rather than a normal curve (i.e. only one peak). Regardless, if more clusters emerge or if the data fits more to a continuum of weight transfer styles, it is likely that different factors will be important for different golfers. This is the focus of study 2.

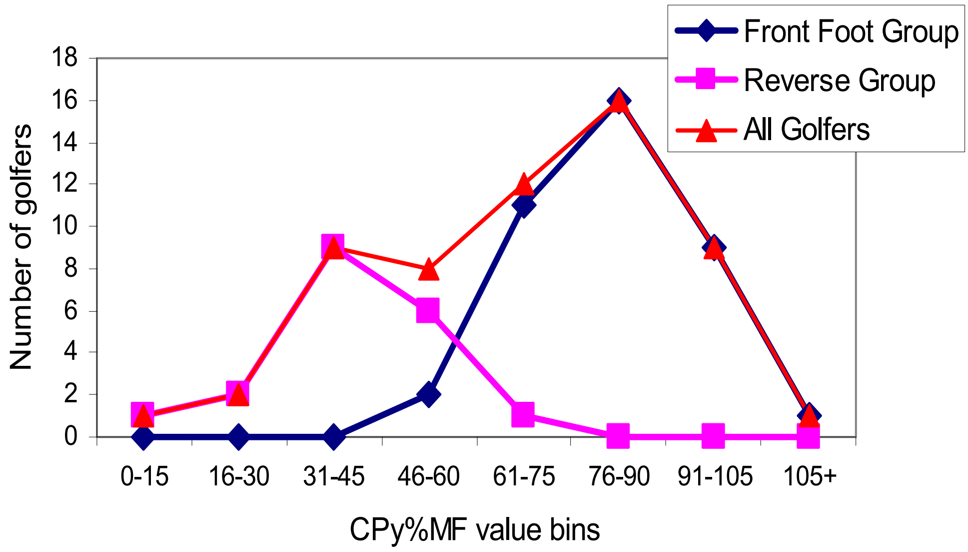


Figure 4.20: Distribution of CPy%MF for All Golfers as well as the Reverse and Front Foot clusters.

4.4.4 'Outlying clusters'

The Extreme Back Foot Reverse group and the Midstance Backswing Front Foot group were considered technical errors rather than useful/valid weight transfer styles. This decision was based on the combined evidence from different sources. First, the groups were only small, with only two golfers in each. Second, the individual cases (golfers) clustered late in the clustering process indicating that their characteristics were considerably different to all other cases in the $N = 62$ data set (and reasonably different to each other). Within-group dissimilarity was further supported by large standard deviations for the Extreme Back Foot Reverse group (TA and MB) and the Midstance backswing Front Foot group (downswing events; table 4.27). Third, the two small clusters both failed at least one of the validation tests (although there are issues with this analysis that will be discussed in the next section). Fourth, and most compelling, the golfers who made up the groups were less skilled (high handicap or social golfers) and/or older (two were over 45) and no low handicap golfers were present in either group. These factors supported the styles as errors or inefficient weight transfer styles, rather than valid styles.

Table 4.27: Standard deviations for CPy% at eight swing events for the four weight transfer groups

Cluster	Label	TA	MB	LB	TB	ED	MD	BC	MF
1 (N=19)	Reverse	5	8	9	13	13	10	12	13
2 (N=2)	Extreme Back Foot Reverse	**18**	**25**	5	1	2	11	6	7
3 (N=39)	Front Foot	5	11	10	9	12	11	10	12
4 (N=2)	Midstance Backswing Front Foot	5	5	0	**20**	**13**	**15**	**13**	**16**

In considering which cluster solution to choose, the 2-cluster solution in which these outlying golfers were part of the Front Foot or Reverse group was considered. In a practical application, this option was possibly the better as some characteristics of each golfer were broadly similar to the major groups and they might simply represent outliers of the large groups rather than separate clusters. That is, the Extreme Back Foot Reverse group is made up of outliers in the Reverse group and the Midstance Backswing Front Foot group is made up of outliers from the Front Foot group. However, as regression analyses were to be performed using each cluster, it was considered more appropriate to treat these golfers as separate clusters rather than part of the larger groups as they were likely to influence the statistical analysis.

4.4.5 Clustering of CPy% and Fz%

Cluster analysis using Fz% was also performed to compare with CPy% data. The same clustering process described in section 4.3.1 was followed but with Fz% at eight swing events used instead of CPy%. Sixty of the 62 golfers reclassified into the same cluster. Of the two golfers that did not reclassify, one moved from the Extreme Back Foot Reverse group to the Reverse group, while the other moved from the Reverse group to the Front Foot group. This golfer showed characteristics of both groups and is discussed more thoroughly in section 4.4.6.2 (golfer 1). Group means for the two large groups were very similar (Figure 4.21). This along with 97% similarity in classification adds more support to the similarity of CPy% and Fz% in the golf swing as well as the underlying cluster structure being valid.

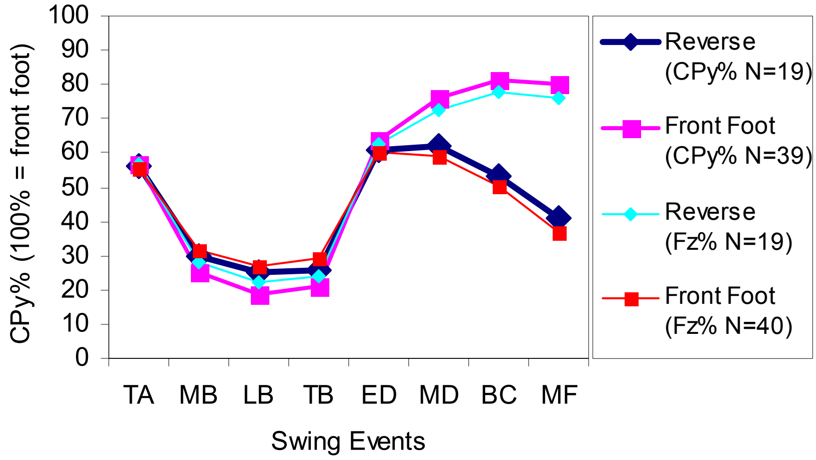


Figure 4.21: Cluster means for the Front Foot and Reverse groups using CPy% and Fz%

4.4.6 Cluster analysis issues

This section discusses some of the cluster analysis issues that were highlighted in the literature review or that were considered in this particular analysis. It should be noted that while it is relatively easy to identify the limitations of each method involved in the cluster process, the positive aspects of the cluster analysis are more difficult to evaluate. There is no 'gold standard' or known underlying cluster structure against which to evaluate the different techniques for real world data (as opposed to artificial data). As such, comparison of different methods and the use of a number of different assessments is the only way to show support for the cluster process used.

4.4.6.1 Number of clusters

In all cluster analyses (original and three replication subsets) evaluating weight transfer swing styles, both the C-Index and Point Biserial Correlation indicated the same solution as optimal in the original analysis and all subset analyses. This provided strong support in each analysis for the particular cluster solution being optimal. It also provided support for the findings of Milligan and Cooper (1985) who reported that these methods were considered strong (C-Index = 3rd; Point Biserial Correlation = 7th of 30 methods). As they were always in agreement, and as the cluster solution indicated was strongly validated, this would support these techniques as being good indicators of the optimal number of clusters in the data.

An interesting aspect of this analysis was that for all analyses (original and three replication subsets), the stepwise method (agglomerative schedule) indicated the same number of clusters as optimal as the Point Biserial Correlation and C-Index calculated on non-hierarchical data. Milligan and Cooper (1985) ranked stepwise method only 11th of the 30 tested and described it as 'mediocre'. However, in this study, the stepwise method performed similarly to the stronger statistical tests used. While generalisation of this finding is limited as the number of analyses performed was low ($N = 4$) and all used the same data set (or subset of the data set), it does highlight differences in performance that might occur with different types of data, as noted by a number of authors (e.g. Everitt, 1979; Milligan and Cooper, 1985; Hair *et al.*, 1995). In the case of this analysis, the use of the agglomerative schedule would have resulted in the same cluster solution (note: the 4-cluster solution was indicated as optimal using hierarchical as well as non-hierarchical data by both C-index and Point Biserial Correlation).

One aspect of the cluster analysis process in this study differed from that used in the literature (e.g. Hair *et al.*, 1995; Milligan and Cooper, 1985). For this study, statistical indices to determine the number of clusters were applied to non-hierarchical data. Other studies have applied the tests on the hierarchical data, determined the optimal cluster solution, and then recalculated only this solution non-hierarchically. In this study, all solutions below a reasonable cut-off (e.g. 11-cluster to 2-cluster solutions) were reanalysed non-hierarchically and then statistical tests were applied to the non-hierarchical cluster solutions to determine which was optimal. This seemed to be a better approach as the non-hierarchical process eliminates nesting by reclassifying cases better and therefore altering the cluster structure. The altered cluster structure could produce different results for the statistical tests and may lead to a different cluster solution indicated as optimal.

To assess if it was more appropriate to evaluate the optimal solution (number of clusters) on hierarchical data or non-hierarchical data, C-Index and Point Biserial Correlation were applied to both. The results of the comparison for Point Biserial Correlation are reported in table 4.28. As can be noted, the optimal cluster solution indicated by each test was the same for hierarchical cluster data and non-hierarchical data. Although not reported here, similar results were evident in all subset analyses and with C-Index. As both processes provided the same result, either could have been used for this data. With the extra analysis required to evaluate all non-hierarchical cluster group means, application of tests to the hierarchical data would have been more efficient. However, the theoretical considerations expressed in the previous paragraph may still hold for different data and application of statistical tests to the non-hierarchical data would seem to be the safer option.

Table 4.28: Point Biserial Correlation coefficients calculated on hierarchical and non-hierarchical cluster data

	11	10	9	8	7	6	5	4	3	2
Hierarchical data	0.48	0.55	0.55	0.55	0.56	0.56	0.5655	**0.5660**	0.55	0.51
Non-Hierarchical data	0.47	0.52	0.52	0.54	0.58	0.58	0.621	**0.622**	0.61	0.58

Another point of note from table 4.28 was the similarity between the 4-cluster and 5-cluster solutions for the Point Biserial Correlation. This indicated that either solution might have been appropriate. However the only difference between the 5-cluster and 4-cluster solutions was that two ‘outlying’ golfers who had remained as clusters of $N = 1$ (i.e. had not clustered with other golfers), combined together to form a cluster of $N = 2$ and did not affect the larger groups in the analysis. As such, either solution would have resulted in the same conclusions as this small group did not pass validity tests.

4.4.6.2 Hierarchical – non-hierarchical process

Point Biserial Correlation results showed stronger effects in the non-hierarchical data in the lower cluster solutions and importantly the optimal solution compared with hierarchical data (refer table 4.28). This indicated that the non-hierarchical process improved the recovery of the underlying cluster structure (i.e. cases were classified more appropriately). This improvement provides support for the use of the hierarchical – non-hierarchical approach recommended by Milligan (1996).

Seven golfers changed clusters from the hierarchical process to the non-hierarchical process. While five of these clustered more appropriately in the final solution, the remaining two golfers provided interesting data that is worthy of highlighting. Figure 4.22 shows the two golfers along with the final group means for the Front Foot group and the Reverse group. Golfer 1 changed from the Front Foot group to the Reverse group and Golfer 2 changed from the Reverse group to the Front Foot group.

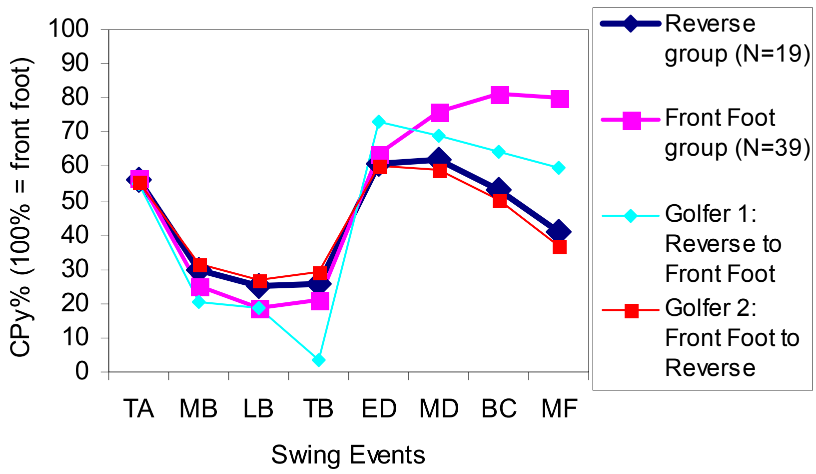


Figure 4.22: Golfers exhibiting unusual weight transfer patterns

The two golfers in figure 4.22 showed similarities with both the Front Foot and Reverse groups. For Golfer 1, while the overall pattern looks more similar to the Reverse group, particularly in downswing, the CPy% magnitudes are more similar to the Front Foot group. This golfer was referred to in section 4.4.5 as clustering in the Front Foot group if Fz% was used compared to the Reverse group when CPy% was used. On a practical level, this golfer might be considered part of the Reverse Group who moves too far forward or a Front Foot golfer who moves forward too early. For Golfer 2, CPy% magnitudes are more similar to the Reverse group in downswing but the pattern looks more similar to the Front Foot group. In practical terms this might suggest that this golfer is a Front Foot golfer that does not shift the weight forward enough (scaling error) or a Reverse golfer who does not rapidly shift the weight forward at ED but who still ends up with weight at the desired (Reverse group) position at BC. While Golfer 2 possessed a high handicap (24), Golfer 1 was a low handicapper (2) so the unusual weight transfer patterns cannot necessarily be discounted as technical errors.

On a practical level, these two golfers would have been difficult to classify for coaching. The analysis might require examination of more CP parameters or the use of kinematic analyses to better determine which style might be most appropriate for each golfer. Individual-based statistical analysis of each golfer's performance would also allow for a more informed decision. However, with only $N = 10$ trials performed, not enough data existed for this evaluation. It should be noted that these golfers represented only 3% of the golfers tested. For 97% of golfers, cluster analysis allocated without ambiguity.

A future direction which might be of use in respect to these golfers is fuzzy clustering. This method of clustering considers each case in terms of percentage membership rather than belonging to one cluster only (e.g. Chau, 2001). For these two golfers, this analysis would have provided information on how much these golfers belonged to each cluster group which could be used to determine how they might be coached, or in scientific terms, how they might be treated in the next stage of analysis (i.e. which style they are allocated to, if any). Certainly, its use will introduce its own problems as more work would be required to evaluate what a 50:50 membership golfer should do – move to one particular style or continue to use some of both styles. However, the allocation of a golfer to one style might also hold problems particularly if they show attributes of both styles and the classification of these golfers into one or the other cluster would be based on marginal differences only. Also, irrespective of the information offered by fuzzy clustering (or any other method of classification) the results in this study using cluster analysis were valid and the method used was appropriate.

4.4.6.3 Measure used in clustering data

Golfer 2 in section 4.4.6.2 also highlighted another issue in cluster analysis – the choice of the measure used to assess differences between cases. While this study used the squared Euclidean distance measure, it may not have been the best measure for classifying golfer 1 and 2 in figure 4.22. For these golfers, more appropriate clustering may have been obtained by using the Pearson’s correlation measure, which clusters cases that are highly correlated (and so would cluster similar patterns rather than similar positions).

To compare the squared Euclidean distance and Pearson’s correlation methods, clusters were reanalysed using both measures (4-cluster solution). Results indicated that both measures produced similar cluster structures. Cluster means differed only slightly, as is evident in figure 4.23, which compares the two large groups for both analyses. Further, 85% of cases clustered into the same groups for each analysis (92% of the major groups reclassified). Of note, from section 4.4.6.2, golfer 1 remained in the Front Foot group but Golfer 2 changed to the Front Foot group from the Reverse group. Both Midstance Backswing Front Foot golfers also moved to the Front Foot group (hence, $N = 60$ is in the 'correl' series in figure 4.23).

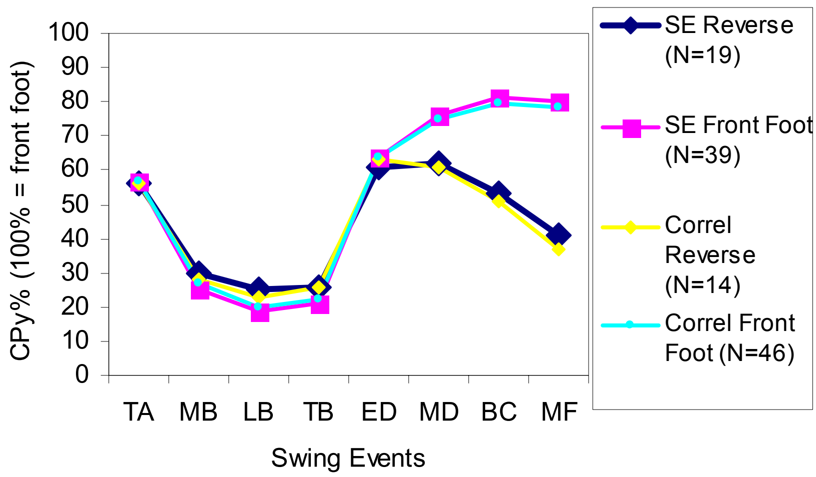


Figure 4.23: Mean CPy% at eight swing events for the Front Foot and Reverse group from cluster analysis using the squared Euclidean distance measure (SE) and Pearson's correlation (Correl) measure

The measure used to cluster golfers in this study was considered prior to analysis. Both the squared Euclidean distance measure and correlation measure had advantages and disadvantages. While the correlation method was better for clustering golfers with similar patterns of weight transfer, it did so with no information on where the pattern existed in relation to the feet. Conversely, the squared Euclidean distance measure provided this information but was less robust for extracting similar patterns. The squared Euclidean distance measure was chosen as the coaching emphasis is on weight position and previous scientific literature had evaluated positions rather than patterns. Regardless, for 85% of cases, either method provided the same results (92%

of the large cluster cases) and group means for both the Front Foot and Reverse groups were very similar. As such, the two weight transfer styles would have been identified using either measure. This also provides further validation for the existence of the two styles, as both appeared using different measures.

4.4.6.4 Validation

4.4.6.4.1 POINT BISERIAL CORRELATION

An interesting result was noted in Point Biserial Correlation validation. While *r*-values varied, and the strongest of these was produced by the 4-cluster solution, all of the Point Biserial Correlations for different cluster solutions were significant at $p < 0.001$ for both the original and replication analyses. This suggested a lack of sensitivity to different cluster structures or a very similar clustering arrangement between cases for all clusters (and therefore similar Point Biserial Correlation results). For example, the change from the 5-cluster to the 4-cluster solution was only two 'outliers' clustering together which might be expected to affect the correlation less than if two large groups were clustered together. The very small change in r-value between these solutions indicates that this is the case. However a small change in *r*-values (0.04) also existed between the 6-cluster and 5-cluster solutions. Given two large groups clustered between the 6-cluster and 5-cluster solutions, it might be expected to show a large change in *r*-value for the Point Biserial Correlation as the cluster solutions were considerably different. As such, the similarity between Point Biserial Correlation values for these two solutions was not due to a small change in clustering. Rather, it indicated that the measure was not sensitive to a large cluster change at a late stage of clustering (i.e. in the last five cluster solutions moving from 62 clusters to 2 clusters). Table 4.29 repeats the Point Biserial Correlation results for original and replication cluster analyses.

Table 4.29: Point Biserial Correlation results for cluster analyses.

	r-value for p=0.001	11	10	9	8	7	6	5	4	3	2
N=62	0.41	0.47	0.52	0.52	0.54	0.58	0.58	0.621	0.622	0.61	0.58
Subset1 (N=41)	0.48					0.56	0.56	0.59	0.66	0.65	0.63
Subset 2 (N=41)	0.48						0.55	0.56	0.57	0.59	0.57
Subset 3 (N=41)	0.48							0.55	0.55	0.61	0.59

This is an important point for this cluster analysis. Point Biserial Correlation would have validated cluster solutions from the 10-cluster solution to the 2-cluster solution based on setting significance at $p < 0.001$ or effect size large ($r > 0.50$) for the original data. Also, all subset cluster solutions would have been validated. This being the case, a poor choice of cluster solution would not have been detected at the validation stage by Point Biserial Correlation (although a good analysis, as has been performed in this study, should not arrive at a poor solution at the validation stage). This highlights the importance of appropriate selection of the number of clusters in a cluster analysis. It also emphasises the need to use more than one technique to validate the solution as recommended by Milligan (1996).

4.4.6.4.2 REPLICATION

Replication subsets indicated that the larger groups (Front Foot and Reverse) were stable. All cases were reclassified into the same cluster as in the original analysis. That is, all Reverse group golfers were clustered together in each subset while all

Front Foot golfers were clustered together in each subset. Also, the patterns of weight transfer for the Front Foot and Reverse groups were evident in each subset, in spite of reduced numbers in each group. If a pattern is unstable, then a reduced *N* forming the group means might be expected to alter the pattern. This was not the case in this analysis for either the Front Foot or the Reverse group.

Replication subsets indicated that neither of the two small groups was stable. The Extreme Back Foot Reverse group appeared in two of three subset analyses while the Midstance backswing Front Foot group appeared in only one subset analysis. This supported other evidence that these groups were not valid; the small size of the cluster, the late clustering (7-cluster and 5-cluster solutions) of the two cases and the fact that they were composed of high handicap or social golfers.

While the results for the small groups can be interpreted as being unstable (and based on the criteria set they should be), it needs to be considered in light of what happens to small groups in replication analysis. While a useful validation method, replication disadvantages small clusters. For example, using the Extreme Back Foot Reverse group ($N = 2$), there was an 89% chance that at least one of the two golfers would be chosen in a sample of two thirds of the original sample [2/3 chance that golfer 1 will be selected + 2/3 chance that golfer 2 will be selected – 4/9 (chance that both are chosen) = 8/9 = 89%]. However there is only 44% chance that both will be chosen. With only $N = 2$ some instability of the mean can be expected, even for valid groups. Hence the removal of one of the two cases can alter the group mean considerably – enough to move the golfer to another cluster and hence not form the original cluster. While this is the strength of the reclassification procedure as more disparate groups are less likely to reclassify, the effect will be more pronounced on smaller groups compared with larger groups and so a limitation of the method exists. While a better method of replication might be bootstrapping to generate a large number of datasets ($N = 1000$ compared with only three in this study), this is not realistic due to the amount of work required. Also, as cluster analysis is heuristic in nature (Milligan, 1996), researcher input is required at different stages of the analysis, limiting automation. It should be noted that if the small cluster was represented in 89% of bootstrap subsets, it would still be outside the likely significance levels that might be set (e.g. appeared in 95% of analyses, $p < 0.05$).

Small groups (i.e. $N = 1$ or 2) might be outliers or might represent valid clusters that represent only a small percentage of the population or that have been under-sampled in the study. As such, a negative replication result should not on its own discount the cluster. Rather, other aspects, such as theoretical assessment of the groups, are also required to assess the validity of small clusters. While the small cluster cannot be considered robust for that particular study if it does not appear in replication analysis, it does still exist. As such, discounting a small cluster needs to be supported by strong theoretical arguments. It was for this reason that replication, other validity tests, clustering issues and theoretical assessment were all considered in making the decision that the small clusters in this study were technical errors rather than a valid and unique technique/styles (see section 4.4.4).

4.4.6.5 Leave-one-out reclassification

Both large groups passed ‘leave-one-out’ reclassification, with 100% success. All Front Foot golfers were reclassified into the Front Foot group and all Reverse golfers were reclassified into the Reverse group. This supported the validity of the two large clusters.

Both the Midstance Backswing Front Foot group and the Extreme Back Foot Reverse group failed ‘leave-one-out’ reclassification. This provided support for these clusters being invalid, or outliers. However, similar limitations exist for the ‘leave-one-out’ reclassification method as for the replication method; large groups tend to be advantaged and small groups tend to be disadvantaged. The removal of one case from a small group is more likely to change the group means more considerably than the removal of one case from a large group. This being the case, it is more likely that small groups will fail the leave-one-out classification procedure. As for replication, the validation results for leave-one-out classification need to be considered along with other indicators as well as a strong theoretical basis for including or discounting the cluster as a valid ad useful technique.

4.5 CONCLUSIONS

The main and most important conclusion of this study is two different styles exist in the weight transfer profile in the golf swing. In this study, two major groups were identified by cluster analysis; named as the Front Foot group and the Reverse group. The Front Foot group began the swing from a balanced position, moved the weight towards the back foot in backswing, rapidly forward in early downswing and continued towards the front foot through ball contact. The Reverse group was similar to the Front Foot group in backswing and early downswing swing events. However from early downswing the forward movement of weight stopped and began to move towards the back foot through ball contact. Both groups included professional or elite amateur golfers and no difference existed in Club Velocity at ball contact or Handicap indicating neither technique was a technical error.

A number of differences existed between the Front Foot and Reverse style of swing. CPy% at swing events near ball contact differed significantly, with the Front Foot group positioning CPy% nearer the front foot in mid downswing (76% compared to 66% for the Reverse group), ball contact (81% compared to 53%) and mid follow through (80% compared to 43%). CPy velocity differed significantly at early downswing (Front Foot = 1.1 $m.s^{-1}$, Reverse = 0.6 $m.s^{-1}$), mid downswing (Front Foot = 0.9 $m.s^{-1}$, Reverse = -0.6 $m.s^{-1}$) and ball contact (Front Foot = 0.5 $m.s^{-1}$, Reverse = -1.2 $m.s^{-1}$). The Front Foot group also achieved a significantly smaller minimum CPy% position (i.e. further towards the back foot, 13% compared to 18%) and a significantly larger maximum CPy% position (i.e. further towards the front foot, 87% compared with 69%). Maximum CPy% position occurred later in the swing for the

Front Foot group (0.01 s after ball contact for the Front Foot group compared to 0.23 s before ball contact for the Reverse group). The Front Foot group also exhibited a greater overall range of weight transfer in absolute and relative terms (0.36 m and 75% compared to 0.27 m and 51%). No difference existed between groups for age, handicap, height or weight.

Cluster analysis has been instrumental in identifying these styles. The cluster solution was validated by a number of different methods. Significant differences existed between styles for internal (CPy% at eight swing events) and external (range, CP velocity at swing events, maximum and minimum CP values) parameters. Point Biserial Correlation was significant at $p < 0.001$ indicating the groups were significantly different. Replication procedures of two-thirds sample re-analysis and leave-one-out classification indicated the solution was robust. A number of important methodological points were evident in the cluster analysis such as the need to use two or more tests to indicate the number of clusters and to validate the solution.

There are a number of useful future directions for this research. Most importantly, this study needs to be repeated with kinematic data to assist in better defining the mechanisms underlying the two techniques identified in this study. The inclusion of anthropometric measures might also identify if the different styles have a basis in different body types. Assessment of a larger number of golfers would be appropriate to identify if more than two styles exist. Finally, the use of other methods such as fuzzy clustering and neural networks could also hold useful information in identifying styles and assessing which style might be appropriate for individual golfers.

CHAPTER 5

STUDY 2

CENTRE OF PRESSURE IN THE GOLF SWING: GROUP-BASED ANALYSIS

5.1 AIMS

5.1.1 General

1. To examine the relationship between weight transfer and performance in the golf swing on a group basis within different weight transfer styles.

5.1.2 Specific

1. To correlate CPy parameters with Club Velocity for the Front Foot group.
2. To determine the most influential CPy parameters in predicting Club Velocity using multiple regression for the Front Foot group.
3. To correlate CPy parameters with Club Velocity for the Reverse group.
4. To determine the most influential CPy parameters in predicting Club Velocity using multiple regression for the Reverse group.

5.2 METHODS

Golfers were analysed within the cluster groups identified in study 1 (the Front Foot group, $N = 39$, and the Reverse group, $N = 19$) and using data obtained in study 1 (CP parameters, Club Velocity, etc.). As the other two smaller clusters were considered to be due to technical errors rather than valid techniques, they are not analysed further here.

5.2.1 Parameters

Parameters used in Study 2 are reported in table 5.1. Figure 5.1 represents the swing events used in this study (provided in study 1 and repeated here).

Table 5.1: Parameters used in Study 2 to assess the relationship between CP and Club Velocity in the golf swing

Performance	
Club Velocity	Immediately before ball contact ($m.s^{-1}$)
Descriptive	
Handicap	
Age	Years (yr)
Height	m
Mass	kg
CP Displacement	
At each swing event	Relative to the distance between the feet (%)
CPy%TA	
CPy%MB	
CPy%LB	
CPy%TB	
CPy%ED	
CPy%MD	
CPy%BC	
CPy%MF	
CPy Velocity	
At each swing event	Instantaneous velocity ($m.s^{-1}$)
VelCPyTA	
VelCPyMB	
VelCPyLB	
VelCPyTB	
VelCPyED	
VelCPyMD	
VelCPyBC	
VelCPyMF	
Other CP parameters	
VMaxCPy	Maximum CPy Velocity ($m.s^{-1}$)
tVMaxCPy	Time of VMaxCPy relative to ball contact (s)
MaxCPy%	Maximum CPy% - furthest position towards front foot (%)
tMaxCPy%	Time of MaxCPy% relative to ball contact (s)
MinCPy%	Minimum CPy% - furthest position towards back foot (%)
tMinCPy%	Time of MinCPy% - relative to ball contact (s)
CPyR	CPy Range in metres (Maximum CPy – Minimum CPy)
CPyR%	CPy Range in % (MaxCPy% - MinCPy%)

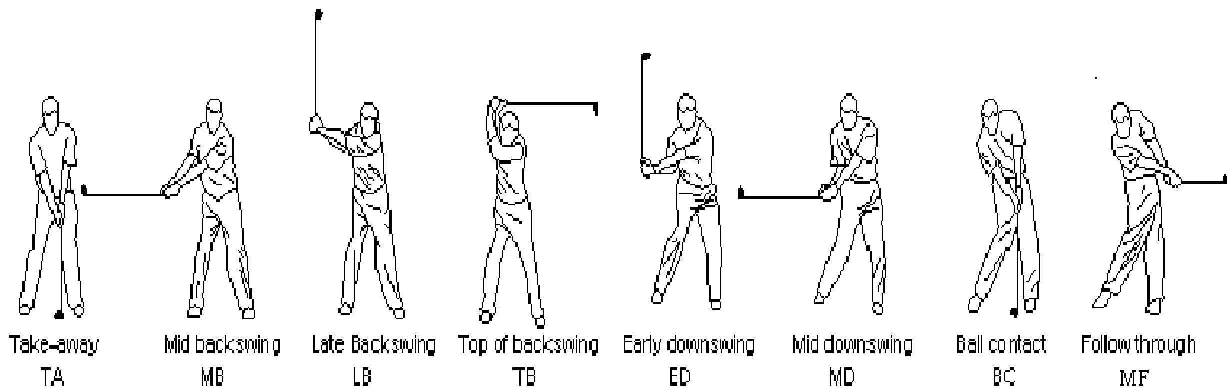


Figure 5.1: Golf swing events used in this study

5.2.2 Statistical analysis

5.2.2.1 Relationship between weight transfer and performance

5.2.2.1.1 CORRELATION ANALYSIS

To examine the relationship between weight transfer (as indicated by CP parameters) and performance (as indicated by club velocity at the instant before ball contact), linear correlations were performed on data within cluster groups (i.e. the Front Foot group and the Reverse group). As Age was found to correlate with Club Velocity within both cluster groups, partial correlations controlling for Age were performed using SPSS 10.0.

Bootstrapping techniques, as outlined by Zhu (1997), were used to establish a 95% confidence level for each correlation. Briefly, this process resamples the existing data to form 1000 datasets with the same N, calculates r-values for each resampled dataset and determines the 25th and 975th ranked r-value to establish a 95% confidence interval. Using the Reverse group Club Velocity - CPyR relationship to illustrate this process, a bootstrap dataset would be formed by randomly selecting $N = 19$ golfers from the original $N = 19$ data. This was performed with replacement which means the selection of each datapoint is made from all $N = 19$ golfers (i.e. the same golfer could be in the dataset more than once). The correlation coefficient between Club Velocity and CPyR was then calculated for the newly formed bootstrap dataset of $N = 19$

golfers. This process was repeated until 1000 datasets were formed and 1000 *r*-values calculated. The correlation coefficients were then sorted in ascending order and the 25th and 975th (2.5% and 97.5%) values were identified to define the 95% confidence limit. The upper and lower limits were considered in analysis of confidence levels to indicate the likely range of correlations that might exist. If these levels crossed zero then the effect was deemed not significant (Hopkins, 1999).

Based on visual inspection of histograms, this researcher was concerned that the assumption of normality may be violated in some parameters. To avoid possible problems with non-normal data, the 'computer intensive method' outlined by Aron and Aron (1999) was used (also used in ANOVA in study 1 and termed "within-dataset distribution" for this study). Using the Reverse group Club Velocity – CPyR relationship as an example, a dataset was formed by randomly selecting a Club Velocity measure from any of the Reverse golfers tested and pairing it with a randomly selected CPyR measure from any of the Reverse golfers tested (e.g. a Club Velocity from Golfer A might be paired with a CPyR from Golfer F). This was repeated (with replacement) until N = 19 data pairs were obtained and a partial correlation controlling for Age was performed on the new dataset. This process was repeated until 1000 datasets (and 1000 *r*-values) were formed. These were sorted in ascending order and the outer 2.5% (25th or 975th) value and 0.5% (5th or 995th) value were determined. These values relate to a two tailed alpha level of p = 0.05 and p = 0.01. If the original *r*-value lay outside of these outer values (i.e. from 1-24 or from 976 to 1000), then this supports the relationship being a true rather than a random effect due to non-normal data. This process differs from the confidence level determination outlined in the previous paragraph in that the parameters are randomly paired together. In the confidence level process, the Club Velocity and CPyR values for any golfer remain together.

Due to the large number of correlations performed, an alpha level of $p = 0.01$ was set to indicate significant analyses. It was felt by this researcher that the $p = 0.01$ level was reasonable balance between reduction of the conventional $p = 0.05$ to avoid type 1 errors (as one in 20 tests could be a chance significant finding) and Bonferroni adjustment, which would have been severe on the alpha level (to gain significance, $p < 0.0001$ would have been required) and increase the chance of a type 2 error. As there was no strong theoretical basis that could be formed from the literature as to which parameters should be chosen (and in fact the increased number of swing events used was to rectify what this researcher saw as a limitation of previous studies), the use of large numbers of parameters was warranted, as there was a need for a greater emphasis on exploratory work. Confidence levels also formed part of the assessment of significance.

Effect sizes were considered to be very important in this analysis, particularly as group sizes were unequal (i.e. a small effect might be significant for the Front Foot group but not the Reverse group). The attention to effect sizes also assisted in stronger interpretation compared to relying on p-values as the only method of identifying important parameters. To assess effect sizes in the correlation analysis, the scale presented by Cohen (1988) was used (table 5.2). The category "Practically No Effect" was added by this researcher to define r-values less than 0.2. This level was not defined by Cohen. Hopkins (2002) suggested a similar scale with more intervals and suggested that $r < 0.2$ could be considered a 'very small effect' and $r < 0.1$ could be considered ‘practically no effect’. As the subject numbers in this study were relatively low, the larger $r = 0.2$ was considered a more appropriate definition for ‘practically no effect’.

Table 5.2: Effect size categories for correlation and regression analysis

r-value	Effect size
Greater than 0.5	Large Effect
0.3 to 0.5	Medium Effect
0.2 to 0.3	Small Effect
Less than 0.2	Practically No Effect

For all analyses, two dimensional (Club Velocity – CP parameter, Age – CP Parameter, Club Velocity – Age) scatterplots were examined to screen for outliers, grouping effects or other abnormalities and to visually assess each relationship. It should be noted that while two dimensional scatterplots were reasonable indicators of the relationships being examined, there is no true graphical representation of partial correlations. Where appropriate, two dimensional scatterplots have been presented in discussion to indicate relationships but the limitations mentioned here need to be considered when viewing.

5.2.2.1.2 MULTIPLE REGRESSION ANALYSIS

5.2.2.1.2.1 Cluster analysis to reduce the number of CP parameters

Due to lower than planned N in this study and the further reduction due to styles being present in the data, the number of parameters used in multiple regression needed to be reduced. To achieve the parameter:case ratio (number of parameters compared with the number of cases or golfers in this study) of 1:5 recommended by Tabachnick and Fidell (1996) as a minimum requirement, cluster analysis was performed (using parameters rather than golfers as in study 1).

Cluster analysis was performed using the same process as described in study 1 but with four differences in this process:

1. Clustering grouped similar parameters together rather than similar golfers.
2. The Pearson's correlation measure was used, rather than the squared Euclidean distance measure, as it operates independently of measurement scales (different scales were used for the different parameters, e.g. CPy%: 1-100, CPy Velocity: approximately -5 – 5 $m.s^{-1}$).
3. Validation of the solution included Point Biserial Correlation as well as cross correlations. Correlations between parameters within and between the cluster groups were visually inspected. A valid solution returned high correlations between parameters within a cluster and low correlations between parameters in different clusters. As the aim of the analysis was to reduce the number of clusters rather than identify a result that was generalisable outside of this study, no replication was performed.
4. The maximum number of clusters was predetermined by the *N* in each group (i.e. seven for the Front Foot group and three for the Reverse group so the 1:5 ratio was achieved). The optimal solution was chosen as the strongest Point Biserial Correlation and C-Index within these constraints.

Once the clusters were formed, one CP parameter was chosen from each cluster for use in regression analysis. This was based on two levels of decision-making:

1. Strongest partial correlation with Club Velocity (controlling for Age).
2. Indicated as significant in previous research (if correlations were similar between parameters and Club Velocity).

Two other methods for parameter reduction were considered along with cluster analysis but were discounted on statistical and theoretical grounds. First, factor analysis was applied to the data but failed diagnostic tests (Kaiser-Meyer-Olkin and Bartlett's test of sphericity < 0.6) indicating it was inappropriate for use with this data. Second, limiting CP parameters to those that were significantly correlated with Club Velocity was also considered and performed in pilot work. However, it was felt that it would be more useful to include parameters that did not necessarily correlate significantly with Club Velocity but might contribute significantly with other parameters in the regression predicting Club Velocity. Given it was the first time styles had been examined and that regression analyses have not been performed on weight transfer data previously, this represented exploratory work. As such, increasing rather than decreasing the number of parameters, within the statistical recommendations for case:parameter ratios, was appropriate.

5.2.2.1.2.2 Best subsets regression

Once the parameters for use in regression analysis were chosen, a best subsets regression was applied to the data (Minitab 11). Best subsets regression uses a combination of Mallow's Cp (total square error) and Best Multiple R^2 assessment, as recommended by Daniel and Wood (1980) to determine the 'best' regression equation for the data. Briefly, a regression equation was calculated for all possible combinations of independent variables (Age was included in all) along with the total square error (Cp) for the regression. The combination of parameters chosen to represent the best subset regression was based on the largest R^2 value for the smallest Cp (error) value. Tabachnick and Fidell (1996) recommend this approach if the parameter:case ratio is low (discussed in section 5.2.2.1.2.1). This process has been used in previous sports biomechanics applications (e.g. Ball *et al*., 2003a; Ball *et al*., 2003b) and is explained in more detail in Results section 5.3.1.2.

5.2.2.1.2.3 Full regression

Once the best subset regression was chosen, a full standard linear regression analysis was performed in SPSS 10. The need to perform both Best Subsets and Full Regression was due to the limited output of data from the Best Subsets regression. While Best Subset identified the best regression for a given set of data and outputs overall R^2 and error values, it does not output information such as change in R^2 for individual parameters in the regression. As such, the use of full regression was not a separate analysis from Best Subsets. Rather it was a repetition of the same analysis using software that would output more information.

As Age was strongly correlated with Club Velocity, it was included in the first block in each regression calculation. The first block refers to the process where Age is entered into the regression first and without any other parameter (to eliminate the effects of Age on the analysis prior to examining CPy – Club Velocity). CP parameters were included in the next block. To assess the importance of individual CPy parameters, change in R^2 and *p*-values for each CPy parameter were examined. The effects of Age on this analysis are discussed in section 5.4.1.4.1.

To examine the robustness of the regression, a subset analysis was performed, where a randomly drawn sub-sample of two-thirds of the original sample (i.e. $N = 28$ for the Front Foot group and $N = 12$ for the Reverse group) was re-analysed. This was an adaptation of the method suggested by Tabachnick and Fidell (1996) where the sample is halved and the regression analysis repeated on both halves. Due to low *N*, analysis of a two-thirds subset, as used by Hodge and Petlichkoff (2000) for cluster analysis and used in study 1, was considered more appropriate for this study. For a parameter (and regression) to be considered robust it should be significant in the subset analysis as well as the original analysis.

5.2.2.1.2.4 Outliers

The data were screened for outliers and influential cases throughout the regression analysis. This was performed on three levels – univariate (z-scores), bivariate (scatterplots) and multivariate (residual analysis and Difference in fit – DFit).

Prior to best subsets analysis, univariate and bivariate outliers were examined. To identify univariate outliers, z-scores were examined within each parameter. A case with a z-score greater than 3.29 ($p < 0.001$; recommended by Tabachnick and Fidell, 1996) was considered an outlier. Bivariate outliers were assessed subjectively from visual inspection of scatterplots (as performed for correlation analysis, section 5.2.2.1.1).

After the best subsets regression and prior to the full standard regression analysis, the selected parameters were included in screening for multivariate outliers using Mahalanobis distance. A level of $p < 0.001$ was set as the cut-off for detecting multivariate outliers as recommended by Tabachnick and Fidell (1996). As this cut-off is different for analyses with different numbers of independent variables, the exact cut-off value is reported in the relevant Results section.

At the completion of the full regression, influential observations (i.e. observations with a notable influence on the R^2 value) were assessed. This was performed using two diagnostic tests. Residuals were examined for each case with a residual value of greater than two considered an outlier (Pedhazur, 1997). DFit, the standardized difference in predicted value with that case removed, was also examined. Cases with a larger DFit influence R^2 values more substantially and need to be examined further (DFit > 1 considered influential, Pedhazur, 1997). This assessment identified cases that either increased or decreased the R^2 result substantially.

Where outliers or influential cases existed, the analysis was performed with these cases removed to examine their influence on the result.

Examples of the use of these screening methods are presented in Results section 5.3.1.2.

5.3 RESULTS

5.3.1 Performance and CP parameters

Table 5.3 reports means and standard deviations for all parameters. Some of these data have already been discussed in Study 1 but are repeated here for easier reference.

Table 5.3: Means and standard deviations for all parameters for the Front Foot group (*N* = 39) and Reverse group (*N* = 19)

	Front Foot group (N = 39)		*Reverse group (N = 19)*	
	Mean	***SD***	***Mean***	***SD***
Performance				
Club Velocity (m.s^{-1})	44.1	3.9	44.1	4.9
Descriptive				
Handicap	11	7	10	9
Age (years)	31.9	12.6	38.1	15.3
Height (m)	1.80	0.06	1.81	0.07
Weight (kg)	80.4	8.6	83.3	10.0
CP Displacement (% between the feet)				
CPy%TA	57	5	56	5
CPy%MB	25	11	30	8
CPy%LB	19	11	25	9
CPy%TB	21	9	26	13
CPy%ED	64	12	61	13
CPy%MD	76	5	62	10
CPy%BC	81	11	53	12
CPy%MF	80	11	41	13
CPy Velocity (m.s^{-1})				
VelCPyTA	0.1	0.3	0.0	0.2
VelCPyMB	-0.3	0.2	-0.4	0.3
VelCPyLB	0.0	0.3	0.0	0.3
VelCPyTB	0.3	0.5	0.1	0.4
VelCPyED	1.1	0.7	0.6	0.9
VelCPyMD	0.9	0.8	-0.6	1.1
VelCPyBC	0.5	0.9	-1.2	1.1
VelCPyMF	-0.3	0.8	-0.3	1.1
Other CP parameters				
VMaxCPy (m.s^{-1})	2.5	0.7	2.5	0.9
tVMaxCPy (s)	-0.14	0.05	-0.18	0.14
MaxCPy% (% between the feet)	87	9	69	9
tMaxCPy% (s)	0.01	0.05	-0.23	0.24
MinCPy% (% between the feet)	13	7	18	8
tMinCPy% (s)	-0.42	0.12	-0.41	0.21
CPyR (m)	0.36	0.07	0.27	0.08
CPyR% (% between the feet)	75	11	51	12

5.3.2 Handicap, Age, Height and Mass

Table 5.4 presents the results of correlations between Club Velocity and Handicap, Age, Height and Mass.

Table 5.4: Correlations between Club Velocity and Handicap, Age, Height and Mass for the Front Foot group (*N* = 39) and Reverse group (*N* = 19).

		Front Foot group (*N* = 39)				Reverse group (*N* = 19)			
		Handicap	Age	Height	Mass	Handicap	Age	Height	Mass
	r	**-0.65**	**-0.59**	0.16	-0.14	**-0.72**	**-0.80**	0.25	0.14
	p	**<0.001**	**<0.001**	0.333	0.39	**0.001**	**<0.001**	0.324	0.569
Confidence limits	2.50%	-0.77	-0.75	-0.12	-0.41	-0.83	-0.88	-0.10	-0.18
	97.50%	-0.53	-0.36	0.39	0.18	-0.60	-0.71	0.55	0.40
Within-dataset distribution	*p*=.05	-0.35	-0.30	0.44	-0.41	-0.50	-0.52	0.55	0.56
	p=.01	-0.44	-0.42	-0.55	-0.50	-0.66	-0.63	0.68	0.66

Large effects were evident between both Handicap and Age with Club Velocity for both the Front Foot and Reverse groups. Both results were in the upper 1% of values based on the within-dataset distribution indicating the significant correlation was not due to issues of non-normal data. While skill level was a desirable trait to include in the analysis, Age was not. Based on this correlation, partial correlations controlling for Age were performed for the remaining analyses. Height and Mass did not correlate with Club Velocity for either group, although a small effect was evident between Club Velocity and Height for the Reverse group. However the level of significance for this result was p = 0.324 only and as confidence interval crossed zero, no significant result was indicated. No outliers existed in these analyses.

5.3.3 CPy% at eight swing events

Table 5.5 presents the results of partial correlations (controlling for age) between Club Velocity and CPy% values at eight swing events.

Table 5.5: Partial correlations (controlling for age) between Club Velocity and CPy% at eight swing events for the Front Foot group (*N* = 39) and Reverse group (*N* = 19).

		CPy%TA	CPy%MB	CPy%LB	CPy%TB	CPy%ED	CPy%MD	CPy%BC	CPy%MF
Front Foot group (N = 39)									
	r	0.07	-0.17	-0.09	-0.05	0.15	0.01	0.12	0.05
	p	0.673	0.316	0.573	0.760	0.363	0.953	0.478	0.757
Confidence Limits	2.50%	-0.16	-0.42	-0.3	-0.29	-0.06	-0.22	-0.14	-0.22
	97.50%	0.28	0.09	0.14	0.22	0.38	0.23	0.37	0.33
Within Dataset Distribution	p=.05	0.46	-0.47	-0.46	-0.45	0.44	0.45	0.44	0.48
	p=.01	0.60	-0.58	-0.61	-0.61	0.54	0.56	0.56	0.62
Reverse Group (N = 19)									
	r	-0.03	0.17	0.44	0.02	0.55	0.28	-0.12	-0.58
	p	0.902	0.524	0.06	0.951	0.023	0.278	0.642	0.015
Confidence Limits	2.50%	-0.26	-0.14	0.12	-0.27	0.34	0.04	-0.37	-0.84
	97.50%	0.16	0.48	0.80	0.46	0.73	0.48	0.15	-0.31
Within-dataset distribution	p=.05	-0.52	0.48	0.49	0.52	0.47	0.46	-0.51	-0.56
	p=.01	-0.67	0.64	0.63	0.62	0.56	0.61	-0.60	-0.70
Reverse Group with outliers removed (all N = 18)									
	r			**0.75**	0.40	0.46			
	p			**0.001**	0.120	0.073			
Confidence Limits	2.50%			0.57	0.18	0.22			
	97.50%			0.92	0.59	0.69			
	N			18	18	18			
Within-dataset distribution	p=.05			0.56	0.48	0.57			
	p=.01			0.73	0.58	0.72			

Bold type = significant at $p < 0.01$

For the Front Foot group, no correlations between Club Velocity and CPy% at eight swing events were significant. There were no r-values greater than 0.2 indicating practically no effect existed for any of the relationships examined. This was supported by all confidence levels crossing zero, indicating that no significant effects existed. No r-values crossed the $p = 0.05$ threshold for within-dataset distribution indicating the analysis was not affected by non-normal data. No outliers existed in these analyses.

For the Reverse group, Club Velocity was not significantly correlated with CPy% at any of the eight events at $p < 0.01$. However, with an outlier ($N = 1$) removed from the partial correlation between CPy%LB and Club Velocity, a large significant effect was evident ($r = 0.75$ with a likely range of 0.57 to 0.92). As well, this result was placed in the upper 1% of the within-dataset distribution indicating the analysis was not influenced by non-normal data. Although not significant, large effects were indicated for CPy%ED and CPy%MF and confidence limits did not cross zero. The large effect between CPy%ED and Club Velocity was due to an outlier, although with it removed, the effect size remained medium and the likely range indicated that there was at least a small effect ($r = 0.22$ to 0.69). CPy%MF was strongly negatively correlated with Club Velocity ($r = -0.58$ with a likely range of -0.31 to -0.84). The *r*-value was placed inside the upper 2.5% of the within-dataset distribution indicating the analysis was not influenced by non-normal data. A medium effect was evident between Club Velocity and CPy%TB with an outlier removed. Also, a small effect was evident between Club Velocity and CPy%MD. However, neither was significant. For the remaining two correlations (Club Velocity and CPy%BC, Club Velocity and CPy%TA) practically no effect existed.

5.3.4 CPy Velocity at eight swing events

Table 5.6 presents the results of partial correlations (controlling for age) between Club Velocity and CPy velocity at eight swing events.

Table 5.6: Partial correlations (controlling for age) between Club Velocity and CPy Velocity at eight swing events for the Front Foot group (*N* = 39) and Reverse group (*N* = 19).

		VelCPyTA	VelCPyMB	VelCPyLB	VelCPyTB	VelCPyED	VelCPyMD	VelCPyBC	VelCPyMF
Front Foot group (*N* = 39)									
	r	-0.27	0.06	0.02	0.06	-0.25	0.24	0.09	-0.11
	p	0.102	0.712	0.908	0.720	0.130	0.150	0.590	0.505
Confidence Limits	2.50%	-0.56	-0.15	-0.26	-0.15	-0.51	-0.04	-0.18	-0.36
	97.50%	-0.08	0.27	0.31	0.28	-0.06	0.51	0.36	0.18
Within Dataset Distribution	*p*=.05	-0.46	0.46	0.47	0.46	-0.46	0.43	0.45	-0.48
	p=.01	-0.59	0.56	0.58	0.59	-0.55	0.58	0.54	-0.62
Reverse Group (*N* = 19)									
	r	-0.42	0.24	0.13	-0.14	-0.26	-0.40	-0.43	-0.35
	p	0.093	0.357	0.608	0.601	0.316	0.113	0.082	0.173
Confidence Limits	2.50%	-0.66	-0.11	-0.1	-0.39	-0.45	-0.63	-0.73	-0.6
	97.50%	-0.2	0.46	0.41	0.07	-0.07	-0.12	-0.15	-0.15
Within-dataset distribution	*p*=.05	-0.52	0.49	0.52	-0.5	-0.49	-0.51	-0.51	-0.53
	p=.01	-0.64	0.56	0.61	-0.59	-0.64	-0.6	-0.66	-0.6
Reverse Group with outliers removed (all *N* = 18)									
	r		-0.03					**-0.69**	
	p		0.922					**0.003**	
Confidence Limits	2.50%		-0.34					-0.87	
	97.50%		0.26					-0.5	
	N		18					18	
Within-dataset distribution	*p*=.05		0.53					-0.61	
	p=.01		0.62					-0.73	

Bold type = significant at $p < 0.01$

For the Front Foot group, no partial correlations between CPy velocity measures and Club Velocity were significant at $p < 0.01$. Small effect sizes existed for VelCPyTA, VelCPyED and VelCPyMD but these were not significant. Further, the confidence limits of VelCPyMD crossed zero, supporting no significant effect. Confidence levels did not cross zero for VelCPyED and VelCPyTA but lower confidence limits were near zero. No r-values crossed the $p = 0.05$ threshold for within-dataset distribution indicating the analysis was not affected by non-normal data. No outliers existed in these analyses.

For the Reverse group, no partial correlations between CPy velocity measures and Club Velocity were significant at $p < 0.01$. However, with an outlier ($N = 1$) removed from the relationship between VelCPyBC and Club Velocity, a large significant effect was returned ($r = -0.69$ with a likely range of –0.50 to –0.87). Within-dataset distribution r-values indicated non-normal data did not influence this result. Although not significant, three correlations returned medium effects for the Reverse group. The relationship between Club Velocity and VelCPyTA returned a medium negative effect with confidence limits indicating a likely range from a small to a large effect and was significant at $p = 0.092$. The medium effects between Club Velocity and VelCPyMD and between Club Velocity and VelCPyMF were not as strong with confidence limits indicating a possible range from large to practically no effect.

A small effect was evident between Club Velocity and VelCPyMB and between Club Velocity and VelCPyED for the Reverse group. However these were not significant and confidence limits crossed zero for VelCPyMB indicating no relationship. As well, the small effect between Club Velocity and VelCPyMB was due to an outlier and when removed, practically no effect existed. Practically no effect existed for the relationships between Club velocity and VelCPyLB and between Club Velocity and VelCPyTB.

5.3.5 Other CPy parameters

Table 5.7 presents the results of partial correlations (controlling for age) between Club Velocity and other selected CPy parameters.

Table 5.7: Partial correlations (controlling for age) between Club Velocity and other CPy parameters for the Front Foot (*N* = 39) and Reverse group (*N* = 19).

		VMaxCPy	tVMaxCPy	MaxCPy%	tMaxCPy%	MinCPy%	tMinCPy%	CPyR	CPyR%
Front Foot group (*N* = 39)									
	r	**0.46**	-0.18	0.15	-0.04	-0.22	-0.12	**0.53**	0.28
	p	**0.004**	0.287	0.367	0.828	0.182	0.47	**0.001**	0.087
Confidence Limits	2.50%	0.20	-0.37	-0.10	-0.35	-0.44	-0.38	0.33	0.04
	97.50%	0.66	0.04	0.37	0.21	0.03	0.18	0.68	0.51
Within Dataset Distribution	*p*=.05	0.33	-0.46	0.47	-0.45	-0.48	-0.47	0.33	0.44
	p=.01	0.41	-0.54	0.56	-0.53	-0.58	-0.67	0.37	0.59
Reverse Group (*N* = 19)									
	r	0.26	0.39	0.3	0.34	0.02	0.25	0.25	0.18
	p	0.31	0.117	0.237	0.182	0.934	0.298	0.31	0.492
Confidence Limits	2.50%	0.10	-0.17	0.06	0.06	-0.28	-0.11	-0.05	-0.11
	97.50%	0.41	0.62	0.56	0.56	0.42	0.51	0.49	0.44
Within-dataset distribution	*p*=.05	0.47	0.45	0.50	0.48	0.50	0.38	0.48	0.47
	p=.01	0.59	0.52	0.62	0.62	0.64	0.52	0.58	0.62
Reverse Group with outliers removed (all *N* = 18)									
	r		-0.17	0.19	0.17	0.39			
	p		0.522	0.47	0.525	0.13			
Confidence Limits	2.50%		-0.48	-0.03	-0.21	0.09			
	97.50%		0.17	0.49	0.52	0.63			
	N		18	18	18	18			
Within-dataset distribution	*p*=.05		-0.55	0.52	0.52	0.55			
	p=.01		-0.68	0.63	0.66	0.69			

Bold type = significant at $p < 0.01$

For the Front Foot group, a large significant effect was evident between Club Velocity and CPyR and a medium significant effect was returned for the relationship between Velocity and VMaxCPy at $p < 0.01$. Both correlations were positive and confidence intervals ranged from medium to large for CPyR and from small to large for

VMaxCPy. As r-values were larger than the thresholds indicated by within-dataset distributions, correlations were not influenced by non-normal data.

A small effect size was evident between Club Velocity and CPyR% (r = 0.28) for the Front Foot group. However, this was not significant and while upper confidence limits suggested the relationship might be large, the lower limits were near zero, suggesting no effect. A small non-significant effect also existed for the relationship between Club Velocity and MinCPy% (r = -0.22). However confidence levels crossed zero indicating no relationship. No significant results were evident for any of the remaining correlations and for all, confidence levels crossed zero. No outliers existed in this data.

For the Reverse group, no correlations between Club Velocity and other CPy parameters were significant at $p < 0.01$, nor were there any large effect sizes. Three correlations returned medium effects (Club Velocity with tVMaxCPy, MaxCPy%, tMaxCPy%). However, all indicated no effect with an outlier removed.

A medium effect size was evident between Club Velocity and MinCPy% with an outlier removed for the Reverse group. However, the effect was not and lower confidence levels indicated practically no effect. Small effects existed between Club Velocity and VMaxCPy, Club Velocity and CPyR and Club Velocity and tMinCPy although these were not significant and lower confidence limits indicated practically no effect. No effect existed between Club Velocity and CPyR%.

5.3.6 Multiple regression analysis: Front Foot group

5.3.6.1 Cluster analysis to reduce the number of parameters

The agglomerative schedule (table 5.8) and visual inspection of the dendrogram indicated that the largest jump in coefficients occurred at the 6-cluster solution for the clustering of CP parameters.

Table 5.8: Selected sections of the agglomerative schedule for hierarchical cluster analysis of Front Foot group cases (*N* = 24 parameters).

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	24	0.872					
2	23	0.816	-0.056	13	12	0.380	-0.056
3	22	0.788	-0.028	14	11	0.375	-0.005
4	21	0.746	-0.042	15	10	0.302	-0.073
5	20	0.722	-0.023	16	9	0.297	-0.005
6	19	0.677	-0.045	17	8	0.189	**-0.108**
7	18	0.625	-0.053	18	7	0.186	-0.003
8	17	0.594	-0.031	19	6	0.061	**-0.125**
9	16	0.588	-0.006	20	5	0.030	-0.031
10	15	0.484	-0.104	21	4	0.022	-0.008
11	14	0.440	-0.044	22	3	0.013	-0.008
12	13	0.436	-0.004	23	2	0.106	-0.093

The 2-cluster to 8-cluster solutions were re-analysed non-hierarchically and statistical tests applied to the non-hierarchical solutions to determine the optimal number of clusters (this process was detailed in study 1). Group means are not reported here. The 6-cluster solution was chosen for further analysis, as it was indicated as optimal by both the C-Index and Point Biserial Correlation (table 5.9).

Table 5.9: Point Biserial Correlation and C-Index data for each solution for the Front Foot group (*N* = 24 parameters). Optimal value for each test in bold.

	8	7	6	5	4	3	2
Point Biserial Correlation	-0.61	-0.62	**-0.63**	-0.57	-0.53	-0.49	-0.39
C Index	1.51	1.62	**1.67**	1.56	1.18	1.17	1.00

Note: For Point Biserial Correlation, the largest negative value is associated with the optimal solution when using the Pearson's correlation measure in cluster analysis.

Table 5.10 reports the CP parameter clusters for the 6-cluster solution. The parameter chosen from the cluster group and the basis for this choice is also reported.

Table 5.10: Cluster groups for the 6-cluster solution for the Front Foot group with the parameter chosen and the basis for the choice.

Cluster		Partial Correlation with Club Velocity	Selected parameter	Basis of selection
1	CPy%TA	**0.07**		
1	VelCPyTB	0.06	CPy%TA	Strongest partial correlation with Club Velocity
2	VelCPyMF	-0.11	VelCPyMF	Only parameter in cluster
3	VelCPyMB	0.06		
3	CPy%MB	-0.17		
3	CPy%LB	-0.09	MinCPy%	Strongest partial correlation with Club Velocity
3	MinCPy%	**-0.22**		
3	tMinCPy	-0.12		
4	CPy%TB	**-0.05**	CPy%TB	Strongest partial correlation with Club Velocity
4	VelCPyLB	0.02		
5	CPy%ED	0.15		
5	CPy%MD	0.01		
5	CPy%BC	0.12		
5	CPy%MF	0.05	CPyR	Strongest partial correlation with Club Velocity
5	VMaxCPy	0.46		
5	MaxCPy%	0.15		
5	CPyR	**0.53**		
5	CPyR%	0.28		
6	VelCPyTA	-0.27		
6	VelCPyED	**-0.25**		Second strongest partial correlation with Club Velocity
6	VelCPyMD	0.24	VelCPyED	
6	VelCPyBC	0.09		Chosen as similar parameter important in previous research
6	tVMaxCPy	-0.18		(Robinson, 1994)
6	tMaxCPy	-0.04		

Cross correlation analysis is presented in table 5.11 with clustered parameters indicted by boxes. All CP parameters with the exception of VelCPyTB produced the strongest bivariate correlation with another CP parameter within the same cluster and not with a CP parameter in another cluster. As well, most parameters were correlated with at least one other parameter within the cluster group at $p < 0.05$ ($r > 0.4$). VelCPyTB correlated more strongly with CPy%MB, which existed in another cluster. However, VelCPyTB would not have been chosen from either cluster as the partial correlation with Club Velocity was lower than other parameters in both clusters.

Table 5.11: Cross correlations between parameters for the Front Foot group (boxes indicate clustered parameters, $r > 0.4$, $p < 0.05$ reported).

	CPy%TA	VelCPyTB	VelCPyMF	VelCPyMB	CPy%MB	CPy%LB	MinCPy%	tMinCPy	CPy%TB	VelCPyLB	CPy%MD	CPy%MB	CPy%BC	CPy%MF	VMaxCPy	MaxCPy%	CPyR	CPyR%	VelCPyTA	VelCPyED	VelCPyMD	VelCPyBC	tVMaxCPy	tMaxCPy
CPy%TA	1																							
VelCPyTB	0.58	1																						
VelCPyMF	.	.	1																					
VelCPyMB	.	0.56	.	1																				
CPy%MB	.	0.66	.	0.72	1																			
CPy%LB	.	0.51	.	0.53	0.66	1																		
MinCPy%	.	.	.	0.65	0.50	.	1																	
tMinCPy	.	.	.	.	.	0.54	.	1																
CPy%TB	.	.	.	.	.	.	.	0.48	1															
VelCPyLB	.	.	.	.	0.48	.	.	.	0.52	1														
CPy%MD	.	.	.	.	.	.	.	.	.	.	1													
CPy%MB	.	.	.	.	.	.	.	.	.	.	0.79	1												
CPy%BC	.	.	0.45	.	.	.	.	.	.	.	0.53	0.71	1											
CPy%MF	.	.	.	.	.	.	.	.	.	.	.	.	0.61	1										
VMaxCPy	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1									
MaxCPy%	.	.	.	.	.	.	.	.	.	.	0.49	0.65	0.87	0.84	.	1								
CPyR	.	.	.	.	.	.	.	.	.	.	.	.	0.55	0.52	0.54	0.63	1							
CPyR%	.	.	.	.	.	0.52	.	.	.	.	.	0.51	0.65	0.66	.	0.74	0.75	1						
VelCPyTA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1					
VelCPyED	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1				
VelCPyMD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1			
VelCPyBC	.	.	.	.	.	.	.	.	.	.	.	.	.	0.50	.	0.42	0.41	.	.	.	0.54	1		
tVMaxCPy	.	.	.	.	.	.	.	.	.	.	.	0.59	.	.	.	.	.	.	.	.	0.68	0.45	1	
tMaxCPy	.	.	.	.	.	.	.	.	.	.	.	.	0.42	.	.	0.58	.	.	.	.	.	0.59	.	1

5.3.6.2 Best subsets regression

Table 5.12 reports univariate z-score data for each golfer in the parameters chosen from cluster analysis to be used in multiple regression analysis for the Front Foot group for univariate outlier assessment.

Table 5.12: Univariate z-score data for parameters chosen from cluster analysis to be used in multiple regression analysis for the Front Foot Group.

Golfer	Club Velocity	Age	CPy%TA	CPy%TB	VelCPyED	VelCPyBC	MinCPy%	CPyR
1	-2.27	2.15	-0.61	1.59	1.15	-0.29	1.73	-1.86
2	-0.20	-0.15	-1.47	1.41	-0.89	1.30	1.39	0.53
3	-2.16	2.15	2.11	-0.05	-0.97	-0.31	-0.75	-1.30
4	0.18	-0.71	-1.72	0.52	-0.30	0.72	0.94	0.11
5	1.52	-0.31	-0.39	-1.84	-2.22	-1.29	-1.31	0.53
6	0.49	-0.95	-0.28	-1.06	1.16	0.80	-0.54	1.38
7	-0.36	-0.55	-0.36	-1.55	-0.63	-0.35	-1.18	-0.87
8	-0.12	-1.34	1.15	-0.70	0.35	0.51	-0.76	0.53
9	1.70	-1.02	-0.39	0.67	0.94	0.97	0.46	0.67
10	-0.40	-0.79	-0.48	-0.06	0.06	-0.38	0.48	-1.01
11	0.85	-1.02	-0.46	0.09	-0.61	-1.17	0.11	-0.59
12	0.53	-0.95	1.52	0.43	-0.28	-0.47	0.83	0.11
13	0.12	0.56	0.41	-0.86	0.82	0.36	-0.44	1.10
14	0.29	1.12	0.03	-1.64	-0.41	1.01	-1.03	1.24
15	0.94	-0.63	-0.36	-0.39	0.86	-0.35	-1.46	0.67
16	-1.13	0.64	1.31	-0.53	2.01	-0.01	0.19	0.82
17	0.01	0.09	0.19	-0.38	-1.10	-1.25	0.30	-1.72
18	-0.92	-0.79	0.49	-0.98	1.49	-1.78	-0.39	-0.03
19	-1.12	1.59	-0.42	-1.30	0.06	-1.14	-1.39	-1.15
20	-0.89	2.15	-0.05	-0.22	0.29	-0.70	0.63	-0.59
21	0.85	1.43	0.57	0.35	0.28	-0.76	-0.20	0.39
22	0.11	0.09	1.06	1.00	-1.16	0.24	-0.38	-1.01
23	-0.49	1.12	-1.11	-0.50	-0.18	0.69	0.41	-0.73
24	0.69	-0.23	0.24	1.24	-1.32	-0.84	-1.88	0.96
25	-0.34	0.01	-1.04	1.31	-2.48	1.32	1.76	-1.30
26	-0.10	0.32	-0.19	-1.29	0.40	1.31	-1.37	0.67
27	0.86	0.01	1.11	1.81	-0.90	0.51	-0.69	1.52
28	-0.15	-0.79	-2.88	0.22	1.25	0.09	0.08	-0.59
29	1.27	-0.71	0.06	-0.07	-0.44	3.06	0.79	1.10
30	-0.38	-1.10	-0.09	0.94	-0.61	-0.27	1.92	-0.73
31	1.61	-0.79	0.80	0.09	0.71	0.22	1.10	1.24
32	0.58	-0.95	-1.05	1.82	0.83	0.19	1.11	-0.59
33	0.80	-0.47	0.16	0.05	-1.14	0.66	-0.16	0.53
34	-1.12	1.27	2.32	1.60	1.11	-0.12	-0.64	0.53
35	-0.30	-0.79	0.36	-0.69	0.71	0.70	-0.39	1.38
36	1.05	-0.71	-0.30	0.15	0.68	0.33	0.59	1.24
37	1.06	-0.15	0.31	-0.53	-0.20	-2.32	0.02	-0.45
38	-1.13	1.04	-0.30	-1.06	0.09	-0.96	-1.25	-1.30
39	-1.93	0.16	-0.27	0.40	0.61	-0.23	1.37	-1.44

No univariate outliers (z-score > 3.29, $p = 0.001$) existed in the data.

Based on visual inspection of two dimensional scatterplots, no outliers existed for bivariate relationships between Club Velocity and the selected CP parameters.

Table 5.13 shows the best subsets regression output from MINITAB software. Likely solutions are in bold type. Figure 5.2 shows the Cp – p graph with the Cp = p line. Repeating for clarity, likely solutions cluster close to this line (Daniel and Wood, 1980).

Table 5.13: Best subsets regression – selected outputs from Minitab software using parameters from cluster analysis for the Front Foot group.

Vars (p)	R^2	Cp	Age	CPy%TA	CPy%TB	VelCPyED	VelCPyMF	MinCPy%	CPyR
2	**53.2**	**2.1**	**X**						**X**
2	39.2	19.5	X			X			
2	38.3	20.3	X					X	
3	**62.3**	**2.6**	**X**			**X**			**X**
3	54.2	8.5	X	X					X
3	53.3	9.4	X					X	X
3	53.2	9.4	X		X				X
3	53.2	9.5	X				X		X
4	63.6	2.4	X	X		X			X
4	62.7	3.2	X			X	X		X
4	**62.3**	**3.6**	**X**			**X**		**X**	**X**
4	**62.3**	**3.6**	**X**		**X**	**X**	**X**		**X**
4	54.5	10.3	X	X				X	X
5	64.0	4.1	X	X		X	X		X
5	63.6	4.4	X	X		X		X	X
5	63.6	4.4	X	X	X	X			X
5	**62.8**	**5.1**	**X**		**X**	**X**	**X**		**X**
5	**62.7**	**5.2**	**X**			**X**	**X**	**X**	**X**
6	**64.1**	**6.0**	**X**	**X**				**X**	**X**
6	**64.1**	**6.0**	**X**	**X**		**X**	**X**	**X**	**X**
6	63.6	6.4	X	X	X	X		X	X
6	62.8	7.1	X		X	X	X	X	X
6	54.8	14.0	X	X	X		X	X	X
7	64.1	8.0	X	X	X	X	X	X	X

*** Likely solutions in bold**

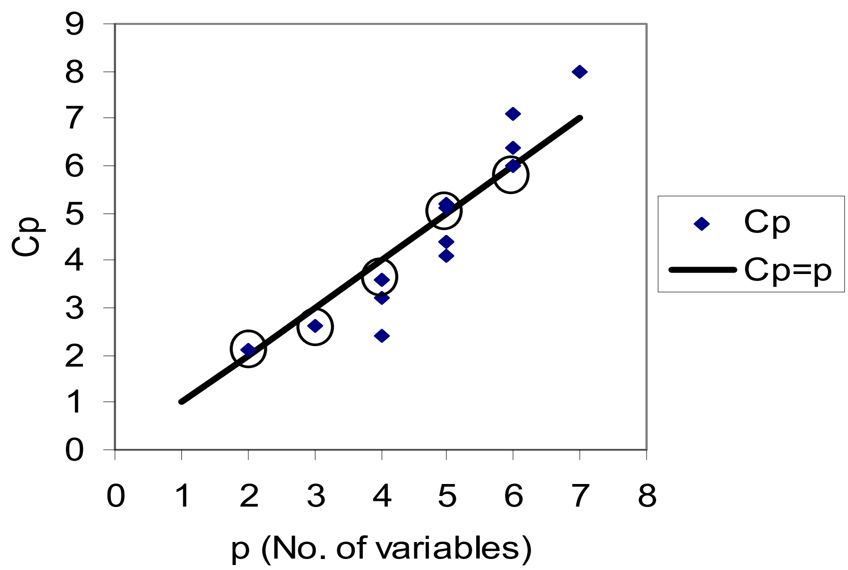


Figure 5.2: Cp-p plot displaying the Cp=p line and likely solutions circled for the Front Foot group (note: vertical axis shortened to better display points close to the Cp=p line. As such, points with Cp > 8 not displayed)

Examining circled points in the Cp-p plot in figure 5.2, 2-parameter, 3-parameter, 4-parameter, 5-parameter and 6-parameter solutions returned Cp values close to the Cp=p line. The 3-variable solution (Age, VelCPyED and CPyR) was chosen to analyse from these possibilities, as it was located near the Cp=p line and the addition of more variables did not increase R^2 values substantially.

5.3.6.3 Full multiple regression analysis using parameters chosen from best subsets for the Front Foot group

To screen for multivariate outliers in the 3-parameter regression, Mahalanobis distance was calculated for each case (Table 5.14).

Table 5.14: Mahalanobis distance for each case using the 3-parameters obtained from cluster and best subsets analysis for the Front Foot Group

Case	M	Case	M	Case	M	Case	M
1	8.04	**11**	2.28	**21**	2.87	**31**	1.97
2	1.41	**12**	0.99	**22**	1.94	**32**	2.9
3	5.85	**13**	2.61	**23**	1.44	**33**	1.96
4	0.59	**14**	4.69	**24**	3.22	**34**	3.48
5	5.97	**15**	1.22	**25**	6.85	**35**	2.29
6	3.11	**16**	4.79	**26**	0.77	**36**	1.88
7	1.75	**17**	3.59	**27**	4.14	**37**	0.25
8	1.94	**18**	3.37	**28**	3.8	**38**	2.13
9	1.97	**19**	2.97	**29**	1.78	**39**	3.01
10	2.68	**20**	4.68	**30**	2.81		

M = Mahalanobis distance. For 3 variables, a Mahalanobis distance > 16.3 (p = 0.001, recommended by Tabachnick and Fidell, 1996) was considered a multivariate outlier.

As no case was indicated as a multivariate outlier, the analysis continued with all cases included. Table 5.15 shows the full regression analysis output for this solution.

Table 5.15: Regression analysis for the 3-parameter solution for the Front Foot group (Age, CPyR, VelCPyED).

	Full Regression	Individual Parameters		
		Age	CPyR	VelCPyED
R^2 / Change in R^2	0.63	0.34	0.22	0.06
p	< 0.001	0.001	< 0.001	0.021
Equation	Club Velocity = 39 - 0.12 Age + 29 CPyR -1.5 VelCPyED			

A strong R^2 value of 63% ($p < 0.001$) was returned for the regression. All parameters were significant at $p < 0.05$. CPyR increased R^2 by 0.22 and VelCPyED increased R^2 by 0.06. Table 5.16 reports residuals and DFIT statistics for each case.

Table 5.16: Residuals and DFIT for each case in the 3-parameter regression for the Front Foot group (from cluster and best subsets analysis). Shaded area indicates possible influential case.

Case	DFIT	Res	Case	DFIT	Res	Case	DFIT	Res	Case	DFIT	Res
1	-0.06	-0.09	**11**	0.29	0.90	**21**	**0.78**	**2.08**	**31**	0.41	1.34
2	-0.36	-1.31	**12**	-0.01	-0.02	**22**	0.14	0.46	**32**	0.42	1.16
3	-0.7	-1.34	**13**	0	0.00	**23**	0.13	0.50	**33**	-0.01	-0.04
4	-0.1	-0.48	**14**	0	0.00	**24**	-0.17	-0.46	**34**	-0.33	-0.85
5	0.34	0.65	**15**	0.25	0.98	**25**	-0.39	-0.68	**35**	-0.59	-1.78
6	-0.15	-0.42	**16**	-0.47	-1.02	**26**	-0.06	-0.26	**36**	0.14	0.48
7	-0.14	-0.49	**17**	0.36	0.90	**27**	-0.13	-0.32	**37**	0.34	1.77
8	-0.42	-1.38	**18**	-0.45	-1.17	**28**	0.16	0.39	**38**	0	-0.01
9	0.6	1.89	**19**	0.08	0.24	**29**	0.12	0.41	**39**	-0.54	-1.46
10	-0.08	-0.25	**20**	0.29	0.63	**30**	-0.34	-0.96			

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)

While all cases passed DFIT tests, Case 21 returned a large residual (standard residual > 2; reported by Minitab software). The regression was re-analysed with this case eliminated to examine its influence. As no multivariate outliers existed in the new data set, as indicated by Mahalanobis distance data (table 5.17), all cases were used in the new regression analysis (table 5.18). DFIT diagnostics were also repeated (table 5.19).

Table 5.17: Mahalanobis distance for each case using the three parameters obtained from cluster and best subsets analysis for the Front Foot Group.

Case	M	Case	M	Case	M	Case	M
1	8.74	**11**	2.18	**21**		**31**	1.93
2	1.37	**12**	0.91	**22**	2.01	**32**	2.80
3	1.02	**13**	2.72	**23**	1.93	**33**	1.84
4	0.51	**14**	5.19	**24**	3.11	**34**	3.86
5	5.77	**15**	1.17	**25**	6.9	**35**	2.25
6	3.11	**16**	4.73	**26**	0.85	**36**	1.84
7	1.65	**17**	3.63	**27**	4.08	**37**	0.22
8	1.98	**18**	3.26	**28**	3.64	**38**	2.48
9	1.98	**19**	3.71	**29**	1.66	**39**	2.87
10	2.54	**20**	5.87	**30**	2.7		

M = Mahalanobis distance. For 3 parameters, a Mahalanobis distance > 16.3 (p = 0.001, recommended by Tabachnick and Fidell, 1996) was considered a multivariate outlier

Table 5.18: Regression analysis for the 3-parameter solution with one influential case removed for the Front Foot group (Age, CPyR, VelCPyED).

	Full Regression	Individual Parameters		
		Age	CPyR	VelCPyED
R^2 / Change in R^2	0.67	0.36	0.22	0.07
P	< 0.001	0.001	0.002	0.010
Equation	Club Velocity = 39.4 - 0.2 Age + 29.9 CPyR – 2.1 VelCPyED			

Table 5.19: Residuals and DFIT for each case in the 3-parameter regression for the Front Foot group (from cluster and best subsets analysis).

Case	DFIT	Res	Case	DFIT	Res	Case	DFIT	Res	Case	DFIT	Res
1	-0.08	0.10	**11**	0.28	0.86	**21**			**31**	0.41	0.50
2	-0.38	-1.32	**12**	-0.01	-0.06	**22**	0.12	0.50	**32**	0.42	0.45
3	-0.65	-1.20	**13**	0.06	0.20	**23**	0.16	0.67	**33**	-0.01	-0.01
4	-0.11	-0.52	**14**	0.06	0.26	**24**	-0.17	-0.40	**34**	-0.33	-0.32
5	0.35	0.75	**15**	0.30	1.07	**25**	-0.54	-0.74	**35**	-0.59	-0.61
6	-0.11	-0.40	**16**	-0.43	-0.89	**26**	-0.04	-0.14	**36**	0.14	0.19
7	-0.18	-0.57	**17**	0.34	0.93	**27**	-0.1	-0.19	**37**	0.34	0.37
8	-0.46	-1.52	**18**	-0.47	-1.27	**28**	0.17	0.36	**38**	0.00	-0.02
9	0.70	2.00	**19**	0.10	0.43	**29**	0.15	0.48	**39**	-0.54	-0.62
10	-0.12	-0.35	**20**	0.39	0.94	**30**	-0.4	-1.13			

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)

With the influential case eliminated, the regression was stronger (R^2 = 0.67, p < 0.001). All included parameters were significant at p < 0.05. VelCPyED showed a slight increase in change in R^2 values (0.07 compared with 0.06) while CPyR remained the same (R^2 = 0.22). No cases failed diagnostic tests (residuals, DFIT, leverage) for the regression with case 21 removed.

5.3.6.3.1 SUBSET REGRESSION RE-ANALYSIS FOR THE FRONT FOOT GROUP

Repeating, a randomly drawn sample of $N = 28$ from the $N = 39$ Front Foot group was formed and the regression was re-analysed. No multivariate outliers existed (table 5.20) so all $N = 28$ golfers were entered into the regression analysis (table 5.21). DFIT for each case is reported in table 5.22.

Table 5.20: Mahalanobis distance for each case for the Front Foot subset Group ($N = 28$).

Case	M	**Case**	M	**Case**	M
1	1.21	**11**	3.05	**21**	2.22
2	1.92	**12**	2.66	**22**	1.98
3	0.78	**13**	1.02	**23**	1.15
4	2.47	**14**	1.34	**24**	1.91
5	0.64	**15**	0.44	**25**	2.85
6	1.88	**16**	4.13	**26**	1.21
7	2.42	**17**	3.31	**27**	0.56
8	1.78	**18**	6.72	**28**	0.13
9	0.84	**19**	1.64		
10	2.10	**20**	1.65		

M = Mahalanobis distance. For 3 parameters, a Mahalanobis distance > 16.3 (p = 0.001, recommended by Tabachnick and Fidell, 1996) was considered a multivariate outlier

Table 5.21: Regression analysis for the 3-parameter solution for the Front Foot subset group (N = 28).

	Full Regression	Individual Parameters		
		Age	CPyR	VelCPyED
R^2 / Change in R^2	0.57	0.28	0.26	0.03
p	< 0.001	0.001	0.002	0.120
Equation	37.3 - 0.11 Age + 33.7 CPyR – 1.1 VelCPyED			

Table 5.22: Residuals and DFIT for each case in the 3-parameter regression for the Front Foot subset group (N = 28).

Case	DFIT	Res	Case	DFIT	Res	Case	DFIT	Res
1	-0.13	-0.60	**11**	0.26	0.85	**21**	-0.51	-1.20
2	0.69	1.04	**12**	-0.01	-0.03	**22**	0.49	1.36
3	-0.29	-0.65	**13**	0.32	0.44	**23**	0.40	0.92
4	-0.24	-0.68	**14**	0.14	0.35	**24**	0.03	0.08
5	-0.58	-1.66	**15**	-0.13	-0.31	**25**	-0.70	-1.20
6	0.70	1.86	**16**	-0.49	-0.61	**26**	0.14	0.42
7	-0.23	-0.56	**17**	-0.12	-0.43	**27**	0.43	1.80
8	0.33	0.87	**18**	-0.08	-0.16	**28**	-0.16	-0.33
9	-0.03	-0.11	**19**	-0.01	-0.01			
10	0.04	0.08	**20**	0.17	0.52			

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)

The Front Foot subset regression was significant. However, while Age and CPyR were significant, VelCPyED was not. It should be noted that based on this group of N = 28 golfers, the change in R^2 due to VelCPyED would not have been significant even with N = 39 (i.e. the non-significant finding for this group of N = 28 golfers was not due to lower N). The regression with VelCPyED removed was still significant (R^2 = 0.54, p < 0.001).

5.3.7 Multiple regression analysis: Reverse group

5.3.7.1 Cluster analysis to reduce the number of parameters for the Reverse group

The agglomerative schedule (table 5.23) and visual inspection of the dendrogram indicated that the largest jump in coefficients occurred at the 2-cluster solution for the clustering of CP parameters.

Table 5.23: Selected sections of the agglomerative schedule for hierarchical cluster analysis of Reverse group cases (*N* = 24 parameters). Largest jumps in bold type.

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	24	7.4					
2	23	8.5	1.11	13	12	21.1	1.26
3	22	10.5	2.03	14	11	22.2	1.08
4	21	10.7	0.14	15	10	24.8	2.58
5	20	10.7	0.02	16	9	25.6	0.82
6	19	14.1	**3.40**	17	8	26.4	0.75
7	18	15.3	1.21	18	7	28.0	1.61
8	17	16.4	1.05	19	6	31.1	**3.11**
9	16	18.4	2.03	20	5	32.8	1.71
10	15	19.3	0.88	21	4	34.0	1.21
11	14	19.5	0.18	22	3	38.9	**4.94**
12	13	19.9	0.41	23	2	52.8	**13.9**

The 2-cluster and 3-cluster solutions were re-analysed non-hierarchically and statistical tests applied to the non-hierarchical solutions to determine the optimal number of clusters (this process was detailed in study 1). While the next two largest jump occurred at the 6th and 19th cluster stage, as the limit of parameters to achieve the 5:1 case to parameter ratio was three, cluster solutions above this threshold were not considered for analysis.

The 3-cluster solution was chosen for further analysis, as it was indicated as optimal by both the C-Index and Point Biserial Correlation (table 5.24).

Table 5.24: Point Biserial Correlation and C-Index data for each solution for the Reverse group (*N* = 24 parameters). Optimal value for each test in bold.

	3	2
Point Biserial Correlation	**-0.49**	-0.39
C-Index	**1.17**	1.00

Note: For Point Biserial Correlation, the largest negative value is associated with the optimal solution when using the Pearson's correlation measure in cluster analysis.

Table 5.25 reports the CP parameter clusters for the 3-cluster solution. The parameter chosen from the cluster group and the basis for this choice is also reported.

Table 5.25: Cluster groups for the 3-cluster solution for the Reverse group with the parameter chosen and the basis for the choice.

Cluster		Partial Correlation with Club Velocity All data (N = 19)	Partial Correlation with Club Velocity Minus outliers (N = 18)	Selected parameter	Basis of selection
1	CPy%TA	-0.03			
1	CPy%MB	0.17	-0.03		
1	CPy%ED	**0.55**			
1	CPy%MD	0.28			
1	VelCPyTA	-0.42			
1	VelCPyTB	-0.14			
1	VelCPyED	-0.26		CPy%ED	Strongest partial correlation with Club Velocity
1	VelCPyMF	-0.35			
1	VMaxCPy	0.26			
1	tVmaxCPy	0.39	-0.17		
1	MaxCPy%	0.30	0.19		
1	tMaxCPy%	0.34	0.17		
1	tMinCPy%	0.44	0.25		
2	CPyR	0.25			
2	CPyR%	0.18			
2	CPy%LB	0.44	**0.75**	CPy%LB	Strongest partial correlation with Club Velocity
2	CPy%TB	0.02			
2	VelCPyMB	0.24			
3	VelCPyLB	0.13			
3	MinCPy%	0.02	0.39		
3	CPy%BC	-0.12		VelCPyBC	Strongest partial correlation with Club Velocity
3	CPy%MF	-0.58			
3	VelCPyMD	-0.40			
3	VelCPyBC	-0.43	**-0.69**		

Cross correlation analysis is presented in table 5.26 with clustered parameters indicated by boxes. Twenty one of the twenty four CP parameters produced their largest bivariate correlation with a parameter in the same cluster. However, three did not. CPy%MB and VelCPyED did not correlate with any parameter within-cluster at $r > 0.4$ but did with a parameter in another cluster. VelCPyTA correlated with one parameter within-cluster but more strongly with five parameters in other clusters.

Table 5.26: Cross correlations between parameters for the Reverse group (boxes indicate clustered parameters, $r > 0.4$, $p < 0.05$ reported except where noted).

	CPy%TA	CPy%MB	CPy%ED	CPy%MD	VelCPyTA	VelCPyTB	VelCPyED	VelCPyMF	VMaxCPy	tVmaxCPy	MaxCPy%	tMaxCPy%	tMinCPy%	CPyR	CPyR%	CPy%LB	CPy%TB	VelCPyMB	VelCPyLB	MinCPy%	CPy%BC	CPy%MF	VelCPyMD	VelCPyBC
CPy%TA	1																							
CPy%MB	.	1																						
CPy%ED	.	.	1																					
CPy%MD	0.54	.	0.55	1																				
VelCPyTA	.	.	.	.	1																			
VelCPyTB	.	.	.	.	.	1																		
VelCPyED	.	.	.	.	.	.	1																	
VelCPyMF	.	.	.	.	.	.	.	1																
VMaxCPy	.	.	0.59	.	.	.	.	.	1															
tVmaxCPy	.	.	.	.	.	.	0.37#	0.45	.	1														
MaxCPy%	0.50	.	0.66	0.76	.	.	.	.	0.44	.	1													
tMaxCPy%	.	.	0.59	.	.	.	.	.	0.52	.	0.43	1												
tMinCPy%	.	0.34*	.	.	.	.	.	.	.	.	.	.	1											
CPyR	.	.	.	.	.	0.56	.	.	0.66	.	0.61	0.42	.	1										
CPyR%	0.41	.	0.40	0.48	0.44	0.64	.	.	0.49	.	0.81	.	.	0.83	1									
CPy%LB	.	0.54	.	.	.	.	.	.	.	.	.	.	.	.	.	1								
CPy%TB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.65	1							
VelCPyMB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.56	0.40	1						
VelCPyLB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.73	.	1					
MinCPy%	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.76	0.79	0.51	0.42	1				
CPy%BC	.	.	.	0.64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1			
CPy%MF	.	.	.	.	0.53	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.46	1		
VelCPyMD	.	.	.	.	.	.	0.40	.	.	.	.	.	.	.	.	.	.	.	.	.	0.70	0.46	1	
VelCPyBC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.71	0.60	1

* largest r-value reported for tMinCPy (also the largest r-value within the classified cluster for CPy%MB)
\# largest r-value within the classified cluster for VelCPyED.

The result that CP parameters can correlate more strongly with a parameter in another cluster, as has occurred with these three parameters, represents a limitation of using the cluster process for grouping parameters. This result can occur at the boundaries of each of the clusters without representing a clustering error. Figure 5.3 is a theoretical example of how a case can be clustered (appropriately) in one cluster but be correlated more strongly with a case in another cluster.

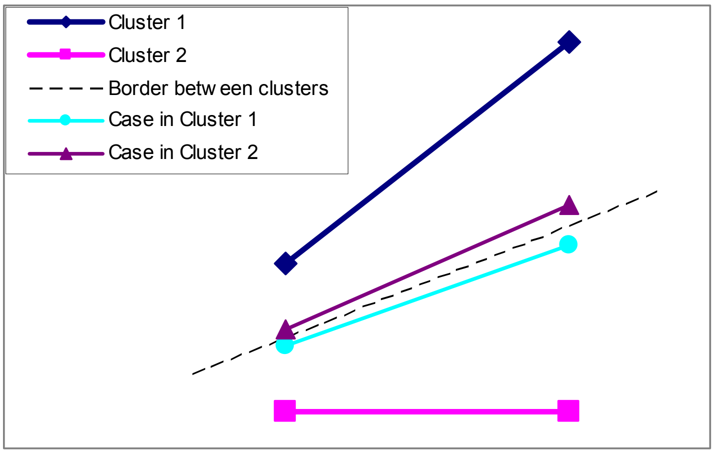


Figure 5.3: Theoretical example of where a case can be more closely associated with a case in another cluster.

A potential problem that this effect can produce is where a case that could be grouped in more than one cluster for this study would be chosen to be entered into regression analysis in one but not the other cluster. For example, CPy%MB was not chosen within the cluster it was grouped with but it might have been chosen from the cluster in which it correlated more strongly with other CP parameters (but was not allocated to). In this case, the clustering process is paramount to the end result. However, this limitation did not affect this analysis. For CPy%MD, VelCPyTA and VelCPyED there were larger partial correlations between Club Velocity and other parameters in all clusters (table 5.27). As such, these parameters would not have been chosen even if they had been classified in a different cluster.

Table 5.27: Correlation coefficients for partial correlations between Club Velocity and CP parameters within each cluster for the Reverse group (*N* = 19 golfers). Bold type indicates strongest partial correlate in each cluster.

Cluster		*r*	*r* (minus outliers)
1	CPy%TA	-0.03	
1	CPy%MB	0.17	-0.03
1	CPy%ED	**0.55**	
1	CPy%MD	0.28	
1	VelCPyTA	-0.42	
1	VelCPyTB	-0.14	
1	VelCPyED	-0.26	
1	VelCPyMF	-0.35	
1	VMaxCPy	0.26	
1	tVmaxCPy	0.39	-0.17
1	MaxCPy%	0.30	0.19
1	tMaxCPy%	0.34	0.17
1	tMinCPy%	0.44	0.25
2	CPyR	0.25	
2	CPyR%	0.18	
2	CPy%LB	0.44	**0.75**
2	CPy%TB	0.02	
2	VelCPyMB	0.24	
3	VelCPyLB	0.13	
3	MinCPy%	0.02	0.39
3	CPy%BC	-0.12	
3	CPy%MF	-0.58	
3	VelCPyMD	-0.40	
3	VelCPyBC	-0.43	**-0.69**

Note: The correlation between Club Velocity and CPy%MF was $r = -0.60$, $p = 0.011$ with the golfer that was an outlier in VelCPyBC removed.

5.3.7.2 Best subsets regression

Table 5.28 reports univariate z-score data for outlier assessment of each golfer for the parameters chosen from cluster analysis for the Reverse group.

Table 5.28: Univariate z-score data for parameters chosen from cluster analysis to be used in multiple regression analysis for the Reverse Group (*N* = 19 golfers).

Golfer	Club Velocity	Age	CPy%ED	CPy%LB	VelCPyBC
1	1.09	-1.25	1.12	-0.47	-1.65
2	0.77	-1.18	-0.37	-0.95	0.74
3	1.25	-1.12	1.18	0.65	-0.24
4	0.50	-1.05	0.42	0.51	-0.65
5	0.81	-1.05	0.08	0.76	-0.15
6	1.44	-0.99	-1.00	1.45	-0.28
7	0.64	-0.92	-0.88	0.63	-0.09
8	-0.04	-0.66	0.48	-0.88	-0.57
9	0.60	-0.20	-0.07	0.20	2.19
10	0.60	-0.13	1.07	-1.00	-1.28
11	-0.79	0.39	-0.82	-1.35	-0.40
12	-1.40	0.72	-2.51	-1.59	1.88
13	-0.51	0.72	1.20	-0.06	0.60
14	-0.03	0.78	1.30	0.76	-1.34
15	-1.40	0.78	-1.16	-1.19	0.76
16	-0.51	0.98	-0.10	0.29	0.40
17	-2.01	1.17	-0.44	1.60	0.72
18	0.20	1.37	0.47	0.43	-0.49
19	-1.20	1.63	0.04	0.72	-0.14

No univariate outliers (z-score > 3.29, $p = 0.001$, as recommended by Tabachnick and Fidell, 1996) existed in the data.

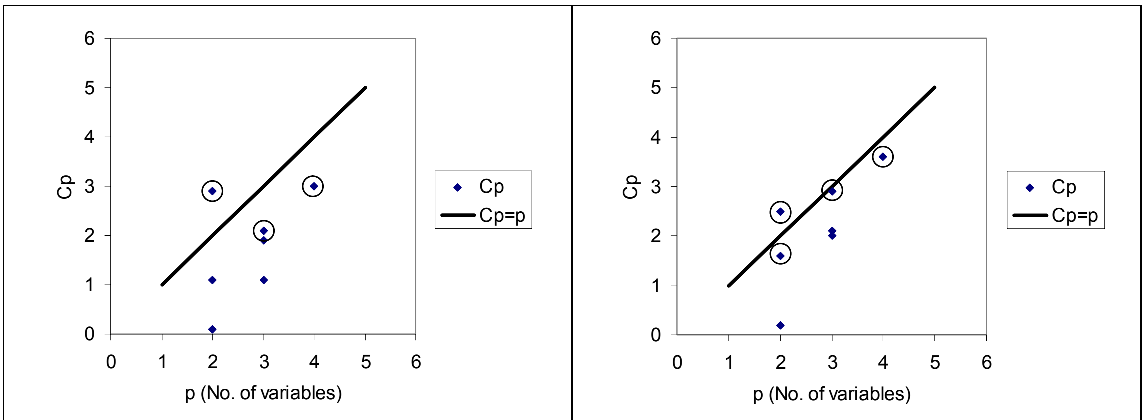
Based on visual inspection of two dimensional scatterplots, two outliers existed for bivariate relationships: one between Club Velocity and CPy%ED and between Club Velocity and CPy%LB (produced by the same golfer) and one between Club Velocity and VelCPyBC. Analysis was performed with and without the two golfers that produced these outliers.

Table 5.29 shows the best subsets regression output from MINITAB software. Likely solutions are in bold type. Figure 5.4 shows the Cp – p graph with the Cp = p line with all data (*N* = 19 golfers) and with outliers removed (*N* = 17 golfers, cases 9 and 17 from table 5.28 were removed). Repeating for clarity, likely solutions cluster close to this line (Daniel and Wood, 1980).

Table 5.29: Best subsets regression – selected outputs from Minitab software using parameters from cluster analysis for the Reverse group.

Vars (p)	R^2	Cp	Age	CPy%LB	CPy%ED	VelCPyBC
All data (*N* = 19 golfers)						
2	**0.70**	**2.9**	x	x		
2	0.69	0.1	x		x	
2	0.66	1.1	x			x
3	**0.73**	**1.1**	x	x	x	
3	**0.72**	**2.1**	x	x		x
3	0.72	1.9	x		x	x
4	**0.74**	**3.0**	x	x	x	x
Minus outliers (*N* = 17 golfers)						
2	**0.76**	**2.5**	x	x		
2	0.75	0.2	x		x	
2	0.75	1.6	x			x
3	0.80	2.0	x	x	x	
3	**0.87**	**2.9**	x	x		x
3	0.75	2.1	x		x	x
4	**0.88**	**3.6**	x	x	x	x

- **Likely solutions in bold**



i. All data (*N* = 19 golfers) | ii. Minus outliers (*N* = 17 golfers)

Figure 5.4: Cp-p plot displaying the Cp=p line and likely solutions circled with and without outliers

The choice of solution of regression for the Reverse group was not as clear as for the Front Foot group. Examining the Reverse group *N* = 19 data, the 2-parameter solution (Age, CPy%LB) and the 3-parameter solution (Age, CPy%LB, VelCPyBC) lay nearest the Cp=p line. While there was a larger R^2 value for another 3-parameter solution (Age, CPy%LB, CPy%ED), the Cp value of 1.1 was considerably further from the Cp=p line. The 4-parameter solution returned the largest R^2 value and a Cp value only slightly further from the Cp=p line. However, there was an increase in R^2 of only 0.2 between the 2-parameter (Age, CPy%LB) and 3-parameter solutions (Age, CPy%LB, VelCPyBC) and the same between the 3-parameter (Age, CPy%LB, VelCPyBC) and 4-parameter solutions.

The choice of solution was clearer with outliers removed (*N* = 17). The 3-parameter solution (Age, CPy%LB, VelCPyBC) produced a considerably larger R^2 value than the 2-parameter solution (Age, CPy%LB) and produced the Cp value nearest the Cp=p line. The 4-parameter solution also produced a Cp value near this line but the increase in R^2 was only 0.1.

It was decided to analyse the 3-parameter solution (Age, CPy%LB, VelCPyBC) as this solution was one of the best with all data (*N* = 19 golfers) and clearly the best with outliers removed (*N* = 17 golfers).

5.3.7.3 *Full multiple regression analysis using parameters chosen from best subsets for the Reverse group*

To screen for multivariate outliers in the 3-parameter regression, Mahalanobis distance was calculated for each case (Table 5.30).

Table 5.30: Mahalanobis distance for each case using the 3-parameters obtained from cluster and best subsets analysis for the Reverse Group

Case	M	Case	M
1	3.79	**11**	4.01
2	3.19	**12**	0.64
3	1.74	**13**	5.87
4	1.75	**14**	2.23
5	1.48	**15**	3.48
6	2.92	**16**	0.98
7	1.35	**17**	4.18
8	1.64	**18**	2.72
9	5.91	**19**	3.47
10	2.82		

M = Mahalanobis distance. For 3 variables, a Mahalanobis distance > 16.3 (p = 0.001, recommended by Tabachnick and Fidell, 1996) was considered a multivariate outlier.

As no case was indicated as a multivariate outlier, the analysis continued with all cases included. Table 5.31 shows the full regression analysis output for this solution.

Table 5.31: Regression analysis for the 3-parameter solution (Age, CPy%LB, VelCPyBC) for the Reverse group (N = 19).

	Full Regression	Individual Parameters		
		Age	CPy%LB	VelCPyBC
R^2 / Change in R^2	0.72	0.64	0.04	0.04
p	< 0.001	< 0.001	0.191	0.186
Equation	48.7 – 0.3 Age + 0.1 CPy%LB – 1.1 VelCPyBC			

A strong R^2 value of 0.72 ($p < 0.001$) was returned for the regression. However, neither CPy%LB nor VelCPyBC were significant. As well, two cases did not pass DFIT and residual analysis (table 5.32).

Table 5.32: Residuals and DFIT for each case in the 3-parameter regression for the Reverse group. Shaded area indicates possible influential cases.

Case	DFIT	Res	**Case**	DFIT	Res
1	-0.09	-0.12	**11**	-0.26	-0.35
2	0.35	0.87	**12**	0.10	0.29
3	0.20	0.43	**13**	-0.27	-0.27
4	-0.11	-0.24	**14**	-0.37	-0.71
5	-0.38	-0.86	**15**	0.16	0.24
6	0.40	0.65	**16**	0.16	0.43
7	-0.12	-0.30	**17**	-2.33	-2.19
8	-0.36	-0.79	**18**	1.21	1.79
9	1.88	1.89	**19**	-0.16	-0.24
10	0.46	0.76			

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)

A second regression analysis was conducted with the two cases that failed DFIT and residual analysis removed (cases 9 and 17). From this regression analysis (*N* = 17 golfers), another two cases failed DFIT and residual analysis. Recalling, Pedhazur (1997) noted that it is common for new cases to become influential in progressive regression analyses. A third regression analysis was performed with the new influential cases removed. From this regression analysis (*N* = 15 golfers), one case failed DFIT and residual analysis. Regression analysis was repeated for a fourth time (*N* = 14 golfers) with the new influential case removed. No case failed diagnostic tests in the fourth regression analysis. This data is summarized in table 5.33 showing regression statistics and diagnostics for each of the four analyses.

Table 5.33: Progressive regression analyses for the Reverse group with influential cases removed until all cases passed diagnostic tests.

	1st Regression (N = 19)			2nd Regression (N = 17)			3rd Regression (N = 15)			4th Regression (N = 14)		
Case	Res	M	DFIT	Res	M	DFIT	Res	M	DFIT	Res	M	DFIT
1	-0.11	3.91	-0.07	0.03	4.33	0.02	0.17	3.85	0.11	1.20	4.69	0.98
2	-0.32	3.19	-0.17	0.49	3.56	0.29	0.48	4.00	0.34	0.55	3.66	0.39
3	0.52	1.68	0.21	0.87	1.66	0.38	1.05	1.35	0.47	1.23	1.20	0.56
4	-0.92	1.49	-0.36	-1.43	1.25	-0.60	-1.76	0.98	-0.79	-1.98	0.90	-0.97
5	-0.23	1.71	-0.09	-0.34	1.81	-0.15	-0.46	1.48	-0.20	-0.66	1.33	-0.29
6	0.78	3.09	0.42	0.94	3.36	0.57	1.15	2.95	0.73	1.15	2.77	0.74
7	-0.30	1.29	-0.11	-0.43	1.42	-0.17	-0.57	1.14	-0.23	-0.77	1.01	-0.32
8	-0.83	1.48	-0.32	-0.93	1.43	-0.39	-1.10	1.22	-0.47	-0.82	1.57	-0.39
9	2.00	5.57	1.70	.	.	.	.	.	.	.	.	.
10	0.90	3.00	0.48	1.47	3.44	0.95	2.04	3.00	1.54	.	.	.
11	-0.39	4.12	-0.24	-0.25	4.06	-0.16	-0.14	3.49	-0.09	0.77	4.33	0.62
12	-0.27	5.57	-0.19	0.39	6.34	0.34	0.39	5.45	0.35	0.27	5.05	0.23
13	0.33	0.71	0.10	0.31	1.20	0.12	0.47	1.85	0.22	0.52	1.66	0.25
14	0.23	3.42	0.13	-0.74	3.62	-0.46	-0.74	6.06	-0.73	-0.21	6.09	-0.21
15	-0.77	2.24	-0.35	-0.99	2.12	-0.48	-1.18	2.02	-0.62	-1.23	1.88	-0.67
16	0.47	1.09	0.16	0.25	1.65	0.10	0.45	3.16	0.28	0.51	2.87	0.31
17	-2.72	4.35	-2.39	.	.	.	.	.	.	.	.	.
18	1.98	2.69	1.15	2.22	2.84	1.50	.	.	.	.	.	.
19	-0.34	3.40	-0.18	-1.70	3.91	-1.22	.	.	.	.	.	.
	R^2/Change in R^2		p	R^2/Change in R^2		p	R^2/Change in R^2		p	R^2/Change in R^2		p
Regression	0.72		<0.001	0.83		<0.001	0.92		<0.001	0.95		<0.001
Age	0.64		<0.001	0.63		<0.001	0.70		<0.001	0.69		<0.001
CPy%LB	0.04		0.191	0.16		0.007	0.16		0.003	0.20		<0.001
VelCPyBC	0.04		0.186	0.07		0.024	0.06		0.014	0.06		0.048

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)
- M (Mahalanobis Distance) > 16.3 considered multivariate outlier (Tabachnick and Fidell,(1996)

While influential cases affected this regression analysis, in all progressive regressions, the R^2 value increased indicating the outliers served to decrease the relationship rather than generate it. As well, once the first outlier was removed, both CPy%LB and VelCPyBC were significant in all regressions where influential cases were removed.

5.3.7.3.1 SUBSET REGRESSION RE-ANALYSIS

A randomly drawn sample of *N* = 13 cases was analysed for the Reverse group. There were no multivariate outliers although there was one bivariate outlier in this subset. Analysis proceeded but the same case did not pass DFIT. A second regression was performed with this case removed, resulting in another case failing DFIT statistics. A third regression was analysed. The analysis was stopped at this point as four cases failed DFIT diagnostics and with these removed, the regression would have been calculated on *N* = 7 cases only. This analysis is summarized in table 5.34.

Table 5.34: Progressive regression analyses for the Reverse group two-thirds replication subset with influential cases removed.

	1st Regression (*N* = 13)			2nd Regression (*N* = 12)			3rd Regression (*N* = 11)		
Case	Res	M	DFIT	Res	M	DFIT	Res	M	DFIT
1	-0.61	0.75	-0.26	-0.17	1.08	-0.08	-0.67	1.12	-0.36
2	-1.17	0.90	-0.56	-1.94	0.75	-1.25	.	.	.
3	-0.52	1.04	-0.24	-0.10	1.27	-0.05	-0.61	1.33	-0.35
4	1.30	2.09	0.95	1.01	2.72	0.90	1.20	2.46	1.16
5	0.15	2.65	0.11	0.57	2.43	0.44	0.25	2.48	0.21
6	0.63	3.14	0.54	-0.23	3.84	-0.25	-0.83	3.71	-1.07
7	0.49	1.01	0.23	1.53	1.17	0.96	1.67	1.25	1.31
8	0.53	8.05	0.85	-0.29	8.21	-0.92	-0.47	7.40	-1.56
9	-0.62	1.80	-0.37	0.17	2.50	0.13	0.24	2.19	0.18
10	-0.44	1.42	-0.23	-0.43	1.29	-0.23	-0.85	1.17	-0.49
11	-0.01	3.57	-0.01	0.21	3.24	0.20	0.36	2.87	0.33
12	-1.24	2.86	-1.10	-0.33	4.49	-0.43	-0.30	4.03	-0.39
13	1.50	6.71	5.49	.	.	.	.	.	.
	R^2/Change in R^2		*p*	R^2/Change in R^2		*p*	R^2/Change in R^2		*p*
Regression	0.79		<0.001	0.93		<0.001	0.97		<0.001
Age	0.54		0.004	0.56		0.005	0.57		0.007
CPy%LB	0.20		0.019	0.19		0.029	0.20		0.028
VelCPyBC	0.05		0.196	0.19		0.020	0.20		0.006

- DFIT (degree of fit) > 1 and/or Res (residual) > 2 considered influential (Pedhazur, 1997)
- M (Mahalanobis Distance) > 16.3 considered multivariate outlier (Tabachnick and Fidell,(1996)

All regressions were significant at $p < 0.05$ in two-thirds subset replication analysis. Examining individual parameters, CPy%LB was significant in all regressions, indicating that it was robust. The result for VelCPyBC was less definite. VelCPyBC was not significant in the first regression but was in the second and third regressions with influential cases removed. The first case removed was also the bivariate outlier identified in correlation analysis for the relationship between VelCPyBC and Club Velocity (refer section 5.3.2.3.2). The most appropriate conclusion from this analysis is that the result for VelCPyBC was also robust as an outlier was the reason for the parameter not being significant, particularly as this outlier was extreme, returning the largest DFIT value of any analysis in this study.

No regression was produced in which all cases passed diagnostics although Pedhazur (1997) reported that it is not uncommon for new cases to fail in progressive regressions and in this event a decision should be made as to how many regressions should be performed. In this analysis, the next regression would have contained too few cases for a reasonable result hence the stopping point. Regardless, for VelCPyBC, the only parameter that was not definitively indicated as robust, R^2 values progressively increased with each progressive regression indicating the influential cases served to reduce its effect rather than produce the significant result.

5.4 DISCUSSION

5.4.1 CPy% at eight swing events

Club Velocity at ball contact was not related to the position of CPy% at the eight swing events used in this study for the Front Foot group. This supported Mason *et al.* (CPy%; 1995) who found no significant association between Club Velocity and CPy% at TA, TB and BC. It also supported Robinson (1994) who found no significant association between Club Velocity and Fz% at TB and BC (but not TA: discussed later). However, comparisons are limited by no study accounting for swing styles prior to analysis. These results indicated that positioning of weight during the swing was not important for Front Foot golfers in developing Club Velocity at ball contact.

In contrast to the Front Foot group, the Reverse group produced significant associations between Club Velocity and CPy% during the swing. The medium effect between CPy%LB and Club Velocity ($r = 0.44$, $p = 0.060$) increased to a large significant effect with an outlier removed ($r = 0.75$, $p = 0.001$). This indicated that positioning weight during the swing was important for the Reverse golfers tested. As well, non-significant medium and large effect were returned for the relationships between Club Velocity and CPy%ED ($r = 0.55$, $p = 0.023$, minus one outlier, $r = 0.46$, $p = 0.073$) and Club Velocity and CPy%MF ($r = -0.58$, $p = 0.015$). With larger subject numbers (e.g. $N = 39$ as for the Front Foot group) these r-values would be significant. This did not support Mason *et al*. (1995) and while Robinson (1994) found an association between Club Velocity and weight position, the swing events of importance were different. However, importantly, neither examined the Reverse style of swing, nor did they examine swing events LB, ED or MF.

A larger CPy%LB, or a CPy position closer to the front foot at late backswing, was associated with a larger Club Velocity at ball contact for the Reverse group. As CPy%LB ranged from 8% to 39%, this result indicated that positioning CPy% nearer 39% of the distance from the back foot to the front foot compared to moving it further towards the back foot (8%) was associated with larger Club Velocities. Handicap and CPy%LB were also related with an outlier removed (r = -0.71, p = 0.001) suggesting more highly skilled golfers positioned CPy% further from the back foot and nearer the midstance position of 50%. However, as Handicap and Club Velocity were themselves correlated (r = -0.72, p = 0.001), this latter finding had limited extra use. There was no association between CPy%LB and Age (r = 0.01, p = 0.902), indicating that the positioning of CPy%LB was not related to age. As mentioned, this swing event has not been used in previous studies so no comparison data exists.

It is possible that the LB swing event is a truer representation when the club accelerates in the direction of the downswing. This being the case, its importance for the Reverse group might be related to obtaining a good weight position further from the back foot at LB from which to generate Club Velocity. Conversely, it might indicate a position too far towards the back foot does not allow for this Club Velocity to be developed as effectively. The top of backswing is the displacement transition point between backswing and downswing and was defined in this study as the instant before the club began downswing. This is also the usual position referred to by coaching texts as the top of backswing (e.g. Leadbetter, 1993). However, the forces required to generate downswing must start before the club begins downswing. It is possible that these forces are initiated nearer the late backswing event. Mean tMinCPy% was 0.43 s before ball contact (65% of the time from TA to BC) which was midway between LB (-0.57 s, 54%) and TB (-0.30 s, 76%) and for all golfers, Fy, the horizontal force in the direction of line of hit (i.e. same direction as CPy%) had changed from negative at TA and MB to positive at LB. Both indicate that the forces producing downswing are initiated nearer LB rather than TB. The start of downswing starting prior to the top of backswing also has support in the literature with Burden *et al.* (1998) noting that maximum hip and shoulder angles were achieved before the top of backswing as defined by club movement. However, as no relationship was indicated for the Front Foot group, the importance of the LB swing event in relation to CPy would seem to be related to the Reverse group only (although other aspects of performance might be important at LB for Front Foot golfers). Future work needs to examine kinematic data to determine other technical aspects of the Reverse swing style compared with Front Foot golfers to better define the mechanism behind the association of CPy%LB and Club Velocity.

Although not significant at $p < 0.01$, large effect sizes were indicated for the relationship between Club Velocity and CPy%ED ($r = 0.55$, $p = 0.023$) and between Club Velocity and CPy%MF ($r = -0.56$, $p = 0.015$) for the Reverse group. A larger CPy%ED, or a CP position closer to the front foot at early downswing, was associated with a larger Club Velocity at ball contact. Considering the ‘reverse’ CPy pattern of Reverse golfers evident from ED to MF, it might be advantageous to move CPy% further to the front foot by the ED event such that there is more distance available for CP to be moved back towards the back foot from ED to MF. This possibility was supported in part by the large negative effect between CPy%ED and the distance CPy% moved between CPy%ED and CPy%BC (i.e. CPy%ED correlated with [CPy%ED - CPy%BC]: $r = -0.76$, $p < 0.001$) indicating a position nearer the front foot at ED was associated with a greater range of movement towards the back foot between ED and BC. CPy%ED was also correlated strongly with VelCPyBC ($r = -0.61$, $p = 0.008$; discussed in section 5.4.2), which indicated a larger CPy%ED, or position nearer the front foot, was related to a larger negative CP velocity at ball contact. As VelCPyBC was more strongly correlated with Club Velocity ($r = -0.69$, $p = 0.003$) than CPy%ED, the position of CPy%ED might be an important transition position to produce a larger negative CP velocity at BC, rather than an important finding in itself. Only a small non-significant effect existed between CPy%ED and CPy%LB (the other CP parameter significantly correlated with Club Velocity for the Reverse group; $r = 0.23$, $p = 0.384$) indicating the association between these parameters and Club Velocity were different.

A smaller CPy%MF, or a position further towards the back foot at mid follow through was associated with a larger Club Velocity at ball contact for the Reverse group. This was a surprising finding as it indicated a weight position nearer the back foot in follow through was advantageous, a finding that is in direct conflict with the coaching literature (e.g. Leadbetter, 1995, Norman, 1996, Grant *et al.*, 1996). Not only was the mean Reverse group CPy%MF value more towards the back foot (40.6%) but golfers who positioned CPy% closer to the back foot produced larger Club Velocities. For example, two golfers currently involved in international amateur and professional tournaments and who achieved some of the larger Club Velocities produced values of 29% and 34%, indicating these golfer positioned CPy well towards the back foot at

MF. Both CPy% and Fz% returned similar correlations with Club Velocity (Club Velocity - CPy%MF, $r = -0.58$, $p = 0.015$; Club Velocity – Fz%MF, $r = -0.57$, $p = 0.021$) showing that the result was not due to horizontal forces (as mentioned, horizontal forces are the only difference between CPy and Fz measures). Not surprisingly, CPy%MF was also related to VelCPyBC ($r = 0.71$, $p = 0.002$) indicating a greater CPy velocity towards the back foot at BC was related to a CPy% position further towards the back foot at MF. Given the proximity of the two events, similar mechanisms are likely to be generating both relationships.

Of importance in this study is that no relationship was evident between CPy% positioning and performance when the data was treated as one group. Partial correlations using all $N = 62$ golfers returned no significant effects and only one small effect existed between Club Velocity and CPy% at eight swing events (Front Foot, Reverse and outliers, $N = 62$, table 5.35). This would have represented a type 1 error for the Reverse group, for which significant effects existed. This result has also highlighted a limitation of previous studies which showed no relationship between position of weight and performance. If the Reverse style existed in these studies, important information would have been masked by the data being treated as one group only with no accounting for styles.

Table 5.35: Partial correlations controlling for Age between CPy% at eight swing events and Club Velocity for all golfers ($N = 62$).

	r	*p*
CPy%TA	0.09	0.486
CPy%MB	-0.05	0.725
CPy%LB	0.11	0.402
CPy%TB	0.05	0.680
CPy%ED	0.19	0.175
CPy%MD	-0.06	0.637
CPy%BC	-0.13	0.321
CPy%MF	-0.23	0.071

The non-significant result in this study between CPy%TA and Club Velocity for both the Front Foot and Reverse groups supported Mason *et al.* (1995) but not Robinson (1994). It was suggested in the literature review of this thesis that the conflict between Mason *et al.* (1995) and Robinson (1994) might have been due to the different range of skill levels tested. Repeating, Mason *et al*. (1995) used golfers from a narrow skill range (single figure handicaps which is likely to reduce ranges of parameters and hence effect sizes) compared with Robinson (1994) who used a wider range of skill levels (professional to high handicap golfers). However, as this study used a similar range of skill level to that used by Robinson and found no relationship between CPy%TA and Club Velocity, this explanation was not supported. Further, reanalyzing data in this study to directly compare with the Robinson study (Fz% between the feet) indicated no relationship existed for the Front Foot group ($r = 0.09$, $p = 0.592$), Reverse group ($r = -0.07$, $p = 0.799$) or with all data included ($N = 62$; $r = 0.05$, $p = 0.728$). As such, this study did not support the finding of Robinson (1994).

Of interest in the non-significant finding for TA in this study for the Front Foot and Reverse groups, as well as that of Mason *et al*. (1995) is that it suggests that the positioning of weight at this event does not influence Club Velocity at ball contact. Further, post-hoc comparison dividing the data in this study into handicap groups showed no difference between golfers of different skill level for CPy%TA (table 5.36). This does not support the coaching emphasis on correct body positioning at TA to assist swing performance (e.g. Grant *et al.,* 1996; Leadbetter, 1995).

Table 5.36: ANOVA comparing different handicap groupings for CPy%TA

Handicap grouping	Front Foot group		Reverse Group	
	F	*p*	*F*	*p*
(1) a. 0-7, b. 8-14, c. 15+ Front Foot *N* (a=13, b=11, c=11) Reverse *N* (a=8, b=5, c=6)	0.43	0.651	0.94	0.356
(2) a. 0-10, b. 20+ Front Foot *N* (a=15, b=6) Reverse *N* (a=10, b=4)	0.21	0.826	2.10	0.150

Note: social golfers eliminated from this analysis
Groupings based on those used in previous studies (1) Koenig *et al*. (1993), (2) Richards *et al*., (1985)
No significant differences between groups for Age at $p < 0.01$

It should be noted that cluster analysis will influence ensuing statistical analysis if the same parameters that were clustered are examined. For this study, cluster analysis grouped similar values of CPy% together to form the Front Foot group and Reverse groups. This reduced the range of CPy% values at each swing event. For example, the range of CPy%BC was 77% for the whole group ($N = 62$) but only 39% for the Front Foot group ($N = 39$). A reduced range of values will reduce effect size and hence statistical power in correlation analysis (Coleman, 1999). As such, statistical power would have been reduced for both groups due to cluster analysis. However, as no effects were evident for the Front Foot group or the group as a whole ($N = 62$), it is unlikely that a type 2 error has been generated by conducting correlation analysis on data that has been clustered. As well, in spite of the decrease in likely ranges and effect sizes, the Reverse group did return strong correlation results for some of the CPy% parameters. In this respect, the use of cluster analysis has actually increased the real power of the study by facilitating the separation of styles from within the data, and allowing for an accurate result to be obtained.

Of particular interest was that CPy% was not related to Club Velocity at TA, TB and BC, key events in coaching and scientific literature. Robinson (1994) expressed surprise at the position of weight at BC being unrelated to performance and not different between amateur and professional players. However, data in this study supported these findings. If this study had used only these events then no significant findings would have been identified. Clearly important information would have been missed for the Reverse group who returned medium and large effects at three swing events not used in previous studies (LB, ED and MF). This finding has provided compelling support for the use of more swing events in the analysis of weight transfer in the golf swing in both scientific and coaching terms.

Based on these findings, CPy% at the eight swing events used in this study holds useful information for the Reverse style of swing but not the Front Foot style. The Reverse group finding of one significant relationship and two relationships returning medium or large effects (although not significant) supports the coaching literature (e.g. Norman, 1995) that the correct positioning of weight during the swing is important to increasing club velocity and hence distance of hit. However, the direction of some of the effects conflicted with the same coaching literature, with a position nearer the back foot at MF, rather than towards the front foot, being advantageous. As well, the Front Foot group findings did not support the coaching literature that weight transfer was important to performance. The finding that different technical aspects are important for the different styles highlights the importance of identifying these styles prior to any statistical or coaching analysis.

5.4.2 CPy Velocity at eight swing events

Club Velocity at ball contact was not related to CPy velocity at the eight swing events used in this study for the Front Foot group. While small effects were evident for three relationships (VelCPyTA, VelCPyED and VelCPyMD) these were not significant. Further, while upper confidence levels indicated these three correlations might have been strong, lower confidence levels indicated no effect (table 5.6, Results section 5.3.4) and large confidence intervals existed within each (*r*-value ranges of 0.5 or 25% of the full scale of correlation values of –1 to 1) which indicated the relationships among the golfers tested were inconsistent. As such, the support for the small effect in these correlations was weak. No data exists in the literature for comparison.

Conversely, for the Reverse group, CPy velocity at the eight swing events used in this study was important. The relationship between Club Velocity and VelCPyBC produced a large effect with an outlier removed. As well, although not significant at $p < 0.01$, medium effects were returned at three other swing events with none of the confidence limits for these relationships crossing zero. This indicated that CPy velocity at swing events provided useful information to the Reverse group. No data exists in the literature for comparison.

The Reverse group produced a medium negative effect between VelCPyBC and Club Velocity at ball contact which was indicated as large with an outlier removed (r = -0.43, p = 0.082; minus one outlier, r = -0.69, p = 0.003). Surprisingly, given VelCPyBC ranged from 1.2 to -2.2 m.s^{-1}, this result indicated that a CP velocity moving more rapidly towards the back foot was associated with a larger Club

Velocity. As can be noted in figure 5.5, only two Reverse golfers produced a positive VelCPyBC value and one of these golfers was the outlier that altered the correlation substantially (denoted as ‘1’). A similar result was obtained using Fz% instead of CPy% with the same two golfers the only ones to produce a positive rate of change of Fz% and the partial correlation between Club Velocity and rate of change of Fz% with an outlier removed for the Reverse group was r = -0.69 (p = 0.003). This indicated that this result was not due to horizontal forces (the only difference between Fz and CP measures). The rapid CPy movement towards the back foot at BC was also relatively large, with VelCPyBC producing the largest magnitude of any swing event calculated (table 5.37). This held when absolute values were used to calculate the mean (i.e. parameters with both positive and negative values will tend to show lower means if absolute values not used). No effect was evident for this parameter for the Front Foot group (r = 0.12, p = 0.487).

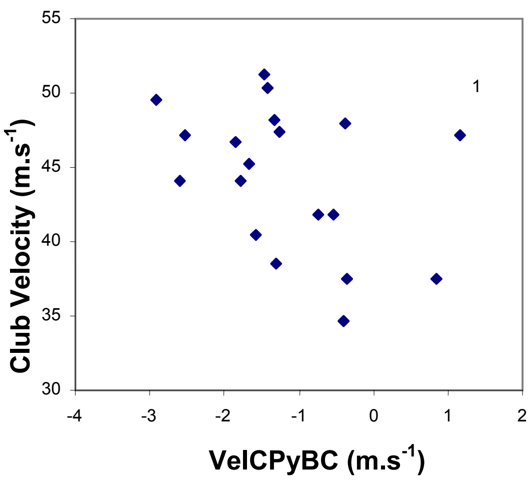


Figure 5.5: Scatterplot of VelCPyBC and Club Velocity (outlier denoted 1). Note that the two dimensional scatterplot is an indication of but not a true representation of the partial correlation analysed in this study.

Table 5.37: Reverse group mean values (*N* = 19) calculated from observed and absolute values for CPy Velocity at eight swing events. All data in m.s^{-1}.

	Mean (observed values)	Mean (absolute values)
VelCPyTA	0.0	0.1
VelCPyMB	-0.4	0.4
VelCPyLB	0.0	0.2
VelCPyTB	0.1	0.3
VelCPyED	0.6	0.9
VelCPyMD	-0.6	0.9
VelCPyBC	-1.2	1.4
VelCPyMF	-0.3	0.9

As discussed in study 1 and in section 5.4.1 of this study, the findings that weight transfer was moving towards the back foot at ball contact was surprising in coaching terms. Even more surprising is the finding in this study that the more rapid movement towards the back foot was associated with a larger Club Velocity for the Reverse group. As such, not only do some golfers shift their weight towards the back foot at ball contact (89% of Reverse golfers produced a negative VelCPyBC), but a more rapid weight transfer was associated with better performance. There is no mention in the coaching literature of weight shift moving towards the back foot at ball contact being a desirable trait. Indeed, this is often associated with technical errors such as a reverse pivot where weight is positioned on the front foot at TB and on the back foot at BC (e.g. Leadbetter, 1993). This finding represents an important change in how weight transfer should be interpreted and coaching practices relating to weight transfer need to consider the implications of this weight transfer style. Not only should a weight transfer moving towards the back foot be encouraged, but a more rapid rate of transfer might be of advantage to the golfer.

The mechanism underlying this finding is not clear from the data and the researcher can only speculate. From visual inspection of video, it was noted that the front knee was extended between ED and BC for Reverse golfers, a technical trait not noted among Front Foot golfers. This is worth evaluating in future exploration of Reverse golfers. If the Reverse golfers are similar to the ‘rotational’ style golfer described by Neal (1998) it might be that the rapid reverse movement of CP assists in positioning the body optimally for rotation. A position nearer midstance would be the better for full body rotation compared with positioning weight nearer the front foot. Rae *et al*. (2001) suggested that weight transfer was not important itself but was needed to position the body for rotation and so was of secondary importance. However, these suggestions are speculative as no data exists to evaluate them but might be useful future directions for analysis of Reverse golfers. More work is required examining the kinematics of the Reverse group to identify the mechanism being utilized and to develop other coaching cues that might be of use in producing and refining this technique.

While no other relationships were statistically significant for either group, relationships that returned small or medium effect sizes with confidence limits that did not cross zero have been briefly discussed below.

For the relationship between Club Velocity and VelCPyTA, a small negative effect was returned by the Front Foot group (r = -0.27, p = 0.102) and a medium negative effect was produced by the Reverse group (r = -0.42, p = 0.09). For both groups, the range of CPy velocities at takeaway crossed zero (Reverse = -0.2 to 0.3 $m.s^{-1}$, Front Foot = –0.6 $m.s^{-1}$ to 0.6 $m.s^{-1}$). As such, this result indicated that a larger negative CPy velocity, or moving CPy more rapidly towards the back foot, was associated with a larger Club Velocity.

Interestingly, 71% of golfers tested produced a positive VelCPyTA value (*N* = 29 Front Foot golfers and *N* = 12 Reverse golfers) indicating that CPy was moving towards the front foot at TA, rather than towards the back foot, for most golfers. This was an interesting finding as it suggested weight was moving forward at the start of the backswing for most golfers. Rate of change of Fz% between the feet was also positive (velocity of Fz%TA = 0.1 $\%.s^{-1}$) so this finding was not due to horizontal forces. This might indicate the 'press' discussed in some coaching (e.g. Grant *et al.*, 1996) and scientific literature (e.g. Wallace *et al*., 1990) where the back leg pushes weight onto the front foot to instigate takeaway. However, the direction of the non-significant effect between Club Velocity and VelCPyTA was negative for both groups indicating moving CPy towards the back foot more rapidly rather than the front foot was associated with a larger Club Velocity. As such, the positive CPy velocity at TA was less advantageous, although the result was not significant.

Medium negative effects existed between Club Velocity and VelCPyMD and between Club Velocity and VelCPyMF for the Reverse group. VelCPyMD was also related to VelCPyBC ($r = 0.60$, $p = 0.018$) so the mechanism generating this result might be similar to that of VelCPyBC. However, VelCPyMF and VelCPyBC were not related ($r = -0.11$, $p = 0.702$). This suggested that not only was a negative CP velocity at BC related to Club Velocity, this relationship also existed in mid follow through and that this result was not simply due to collinearity. However, neither relationship was significant. As well, while confidence levels indicated effects might be large the lower confidence limit suggested no effect might also exist. As such, there is only weak evidence to support any substantial effect existing for these parameters for the Reverse group. No effect existed for VelCPyMF for the Front Foot group and while a small effect existed for VelCPyMD this was not significant and confidence limits crossed zero indicating no relationship.

5.4.3 Other CPy parameters

A larger absolute range of CPy (CPyR: $r = 0.53$, $p = 0.001$) and larger maximum CPy velocity (VMaxCPy: $r = 0.46$, $p = 0.003$) was associated with a larger Club Velocity for the Front Foot group. This suggested that weight transfer was important in the golf swing for the Front Foot group. This supported Koenig *et al.* (1993) and Wallace *et al.* (1990), who both suggested (without statistical evidence) that these factors were important in the golf swing based on low handicap golfers producing larger range and rate of weight transfer compared with high handicap golfers. Importantly, this study is the first to provide group-based statistical evidence that these relationships exist.

CPyR and VMaxCPy were themselves significantly correlated for the Front Foot group ($r = 0.54$, $p < 0.001$). It is possible that the mechanism behind these significant results might also be related. It may be that this mechanism is part of the kinetic chain or proximal to distal sequencing found in other striking activities (e.g. Tennis serve: Elliot *et al*., 1986; Kicking: Putnam, 1993; general: Kreighbaum and Barthels, 1985). That is, the greater velocity of weight transfer, as indicated by VMaxCPy (whole body or proximal movement) develops greater system momentum, which can then be transferred to the club and ball (distal movement). A larger weight transfer range may facilitate this by allowing greater relative distance over which this velocity and momentum can be generated (or be a product of it). This would support Leadbetter

(1995) who suggested weight transfer was essential to developing momentum in the swing. However, discussion of this mechanism is limited by the absence of kinematic data. This is a useful area for future work when kinetic and kinematic data can be combined.

While a strong effect existed between Club Velocity and CPyR for the Front Foot group, only a small association was evident between Club Velocity and CPyR%. This indicated that increasing the CPy range in metres was the more important factor, rather increasing the range of CPy% between the feet. Stance Width (distance between the feet) was also correlated with Club Velocity (r = 0.47, p = 0.005) and with CPyR (r = 0.51, p = 0.001). A similar pattern existed for a subset of Front Foot golfers who were under 40 years of age (N = 29: Stance Width and Club Velocity: r = 0.51, p = 0.005; Stance Width and CPyR: r = 0.44, p = 0.017: subset discussed in section 5.4.1.5.1). In an applied sense, while the cue to increase weight transfer range or increase rate of weight transfer might be useful, a simpler cue of increase stance width may also develop the technical change required to increase CPyR. Leadbetter (1993) suggested that an increase in stance width does not lead to greater distance. Given that club velocity at ball contact is related to distance, the findings in this study are in conflict with these comments.

In contrast to the findings for the Front Foot group, CPyR and VMaxCPy were not related to Club Velocity for the Reverse group. While a small positive effect was evident for both, neither was significant and confidence limits crossed zero for CPyR (Club Velocity – VMaxCPy; r = 0.26 with a likely range of 0.10 to 0.41, p = 0.310; Club Velocity – CPyR; r = 0.25 with a likely range of -0.11 to 0.44, p = 0.318).

Interestingly, mean VMaxCPy values were similar (Front Foot group = 2.5 ± 0.7 $m.s^{-1}$; Reverse group = 2.5 ± 0.9 $m.s^{-1}$). However, the importance of this parameter within each group was different (VMaxCPy correlation with Club Velocity: Front Foot group, $r = 0.46$, $p = 0.001$; Reverse group, $r = 0.26$, $p = 0.310$). This suggested a technical similarity between each group but a distinctly different influence of this technical element on performance. In the case of CPyR, neither the value nor the correlation with Club Velocity was similar. The small value of CPyR and small effect for the relationship between CPyR and Club Velocity (mean = 0.27 m, $r = 0.25$, $p = 0.310$) compared with the significantly larger value and large effect for the Front Foot group (mean = 0.36 m, $r = 0.53$, $p = 0.001$) indicating a different technical element as well as different influence on performance of CPyR. Further, mean Stance Width was similar between groups (Front Foot = 0.51 m; Reverse = 0.52 m) but for the Reverse group did not correlate with Club Velocity ($r = 0.21$, $p = 0.416$). These results provided further support for the Front Foot and Reverse styles being distinct techniques.

While VMaxCPy and Club Velocity were associated for the Front Foot group, the time at which the maximum occurred was not ($r = -0.18$, $p = 0.29$). This suggested that while rapid weight transfer was important, the time at which it occurred was not. Examination of the data indicated that 87% of Front Foot group golfers ($N = 34$) achieved the maximum between TB and ED and so as the range of values for most golfers was small, this reduced statistical power. Similarly, there was practically no effect between Club Velocity and the other timing measures used in this study for the Front Foot group (tMaxCPy, $r = -0.04$, $p = 0.828$; tMinCPy%, $r = -0.12$, $p = 0.470$).

The relationships between MaxCPy% and Club Velocity and MinCPy% and Club Velocity were not significant for either the Front Foot group or the Reverse group. While a small effect existed for MinCPy% for the Front Foot group, confidence levels crossed zero indicating no relationship existed. For the Reverse group, while MaxCPy returned a medium effect size with Club Velocity, this was due to the influence of an outlier ($r = 0.30$, $p = 0.237$; with outlier removed $r = 0.19$, $p = 0.470$). MinCPy returned a medium effect size with an outlier removed ($r = 0.02$, $p = 0.934$; with outlier removed, $r = 0.39$, $p = 0.130$), although this was not significant. However, it would be worth reevaluating MinCPy% for the Reverse style in future work with a larger N as $r = 0.39$ would be significant at $p < 0.01$ with $N = 40$ golfers and at $p < 0.05$ with $N = 29$ golfers.

There was little support for the importance of timing parameters for the Reverse group. An outlier produced the non-significant medium effect sizes evident between Club Velocity and time-based CP parameters (tVMaxCPy, tMaxCPy%, tMinCPy%) and with the outlier removed, no association was indicated. For the relationships between Club Velocity and tVMaxCPy, and between Club Velocity and tMaxCPy%, practically no effect existed without the outlier. A small effect remained for the relationship between Club Velocity and tMinCPy%, although the level of significance for this result was only $p = 0.330$. As well, for all correlations with outliers removed, confidence levels crossed zero, indicating no effect existed. These results indicate that the timing parameters measured in this study were not related to Club Velocity for the Reverse group.

It is surprising given the coaching emphasis that exists on tempo and timing of the swing that no timing parameters were related to Club Velocity. Further, there was no relationship between any of the timing parameters and handicap, or between handicap groups (table 5.38). This analysis was limited to the Front Foot group due to small N for the Reverse group combined with the large influence of outliers. It might be that the golfers in the group had optimized their own swing so the effects of CPy timing were not evident or that differences are not evident on a group basis. Intra-individual analysis may also be required to define the importance of weight transfer timing parameters.

Table 5.38: Post Hoc statistical analysis of CPy timing measures (Front Foot group only).

Handicap grouping	tVMaxCPy		tMaxCPy%		tMinCPy%	
	F / r	*p*	*F / r*	*p*	*F / r*	*p*
1. 0-7, 8-14, 15+ *N* (a=13, b=11, c=11)	0.46	0.699	0.27	0.766	0.002	0.998
2. 0-10, 20+ *N* (a=15, b=6)	1.09	0.744	0.42	0.525	0.08	0.776
Partial Correlation with Handicap controlling for Age (*N* = 39)	-0.10	0.570	0.14	0.444	0.01	0.973

Note: social golfers eliminated from this analysis
Groupings based on those used in previous studies 1. Koenig *et al*. (1993), 2. Richards *et al*., (1985)

Examination of scatterplots of individual golfer data indicated a measurement issue might have influenced the examination of timing data. For five golfers, maximum CPy values appeared in distinct clusters. A similar outcome was evident in minimum CPy for three golfers. Figure 5.6 shows two datapoint that are separated from the remaining eight points for a selected golfer for MaxCPy%. These clusters were associated with either the first or second peaks denoted by arrows in figure 5.7.

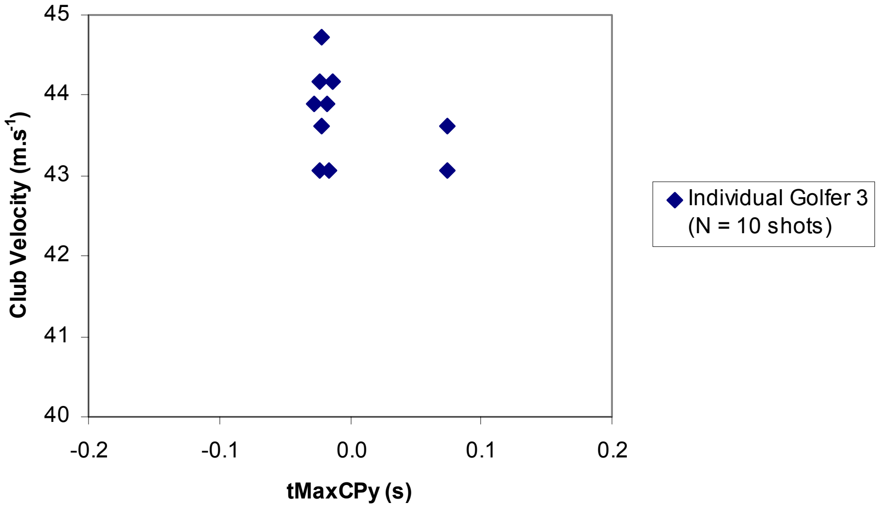


Figure 5.6: Scatterplot of Club Velocity and tMaxCPy for a selected golfer (*N* = 10 shots), showing two clusters of points due to maximum CPy occurring at distinctly different points in the swing

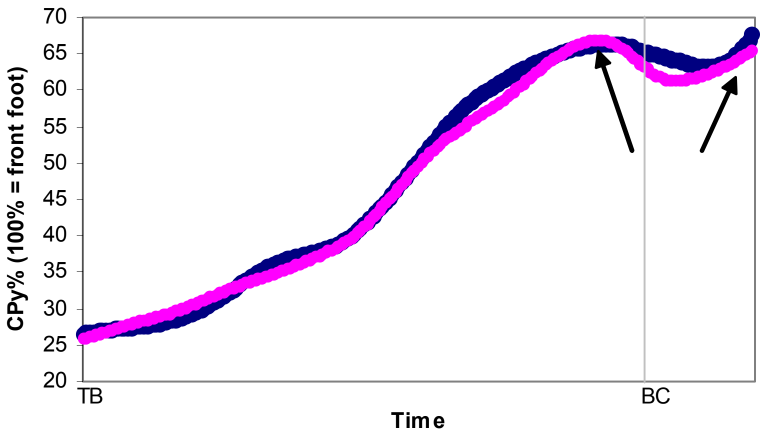


Figure 5.7: CPy%-time curve for a golfer who produced clusters in tMaxCPy showing two different points at which the maximum occurred (denoted by arrows).

Two problems are produced from this issue. First, the use of the mean time value is not representative of either peak. Second, the peaks relate to very different parts of the swing. Arguments exist for leaving this data intact as well as for eliminating it from the analysis. The data is, by definition, the maximum CPy value between TA and MF and therefore it can be reasoned that it should remain. Alternatively, as the mean value across $N = 10$ shots was used in further evaluation, this value could be considered a misrepresentation of what was intended to be measured. In the case of figure 5.7, the maximum occurred at a distinctly different time in the swing and existed on another ‘hill’ in the CPy curve. The mean value for the time the maximum occurred will not represent the region of either maximum (i.e. near each of the arrows) and would indicate maximum occurred near the trough just after BC. Further, although this was not the case with all golfers, the CPy% maximum occurred at the measurement limit and given the upward direction of the curve at this point might be expected to further increase. As such, it might not represent a maximum at all. It should be noted that this issue was not evident in golfers used to evaluate stability of the mean otherwise it would have been identified at that point.

Re-evaluation of this measurement issue indicated it did not affect the results of statistical analyses in this study. Correlations between Club Velocity and potentially affected CP parameters were re-evaluated in two ways with no analysis changed the significant and non-significant results. In the first re-evaluation method, the golfers who produced clusters in timing data were removed from the analysis (three for tMinCPy, five for tMaxCPy). In the second, these golfers were included but maximum CPy% was re-calculated for these golfers between TA and BC (rather than TA to MF). This produced one cluster of results for all but one golfer (this golfer was not included in this analysis and is discussed in detail in the next study). MaxCPy%, MinCPy%, CPyR and CPyR% were also re-evaluated as they might also have been compromised. As indicated in table 5.39, while some r-values were altered, none became significant (or not significant). As such, while future research needs to be mindful of this measurement issue, it did not affect the decisions made in this study.

Table 5.39: Partial correlations between Club Velocity and CP parameters possibly affected by the timing measurement issue

	Original Analysis		All but with mean recalculated for problem golfers		Problem golfers eliminated		
	r	*p*	*r*	*p*	*r*	*p*	*N*
Reverse	*N* = 18		*N* = 17		*N* as indicated		
MaxCPy%	0.30	0.237	0.31	0.230	0.26	0.410	15
tMaxCPy%	0.34	0.182	0.36	0.161	0.41	0.126	15
MinCPy%	0.02	0.934	-0.15	0.554	-0.02	0.941	16
tMinCPy%	0.25	0.298	0.32	0.217	0.38	0.218	16
CPyR	0.25	0.310	0.35	0.189	0.35	0.265	14
CPyR%	0.18	0.492	0.10	0.708	0.21	0.508	14
Front Foot	*N* = 39		*N* = 39				
MaxCPy%	0.15	0.367	0.04	0.816	-0.02	0.928	37
tMaxCPy%	-0.04	0.828	0.11	0.528	-0.19	0.262	37
MinCPy%	-0.22	0.182	-0.22	0.175	-0.21	0.219	38
tMinCPy%	-0.12	0.470	-0.07	0.680	-0.07	0.679	38
CPyR	**0.53**	**0.001**	**0.47**	**0.003**	**0.48**	**0.003**	**36**
CPyR%	0.28	0.087	0.28	0.089	0.28	0.091	36

Note: The outlier identified in the Reverse group analysis (section 5.3.5) was eliminated for all datasets in this table

5.4.4 Multiple regression analysis

The combination of CPyR and VelCPyED predicted 28% of the variance in Club Velocity for the Front Foot group. CPyR was the major weight transfer influence, accounting for 20% of this 28% with VelCPyED contributing 8%. Two-thirds sub-sample analysis indicated that the result was robust for CPyR but not for VelCPyED, as it was not significant in the reanalyzed regression. So while both were included, the importance of VelCPyED is questionable and would require more research to substantiate. While Robinson (1994) conducted regression analyses, these included kinematic variables and so comparison is not possible. No other study has conducted multiple regression analysis on weight transfer parameters.

As noted in the previous section, CPyR and VMaxCPy were related for the Front Foot group and hence did not both appear in the best regression due to shared variance. However, it is important not to discount VMaxCPy on this basis as less important as r-values varied only slightly (VMaxCPy: $r = 0.46$ compared with CPyR: $r = 0.53$). Hair *et al.* (1995) also warned against this type of interpretation when collinearity exists between parameters. An example of the potential risk of considering a parameter unimportant if not entered into a regression was highlighted when analyzing a subset of the data (golfers under 40 years old) in this study. In this analysis VMaxCPy was included in the best subset regression and returned a larger r-value for the correlation with Club Velocity (VMaxCPy: $r = 0.61$; CPyR: $r = 0.53$). Further discussion of the subset group is in section 5.4.4.1.1.

For the Reverse group, both significant correlates (CPy%LB and VelCPyBC) appeared in the multiple regression analysis predicting Club Velocity from CP parameters. With outliers removed, CPy%LB produced a change in R^2 of 16%, while VelCPyBC produced a change in R^2 of 7%. Both were significant in sub-sample regressions, indicating that the association between these parameters and Club Velocity was robust for the Reverse group. As both appeared in the regression, this suggested that both parameters were associated with Club Velocity at ball contact independent of each other. This was supported by the correlation between the parameters indicating practically no effect ($r = -0.01$, $p = 0.988$). The combined total of 23% of the variance in Club Velocity accounted for by CP parameters was slightly smaller than the 28% accounted for in the Front Foot group.

For the Front Foot group, while the result was not robust, it is appropriate to examine VelCPyED in more detail as it appeared in the best regression. From correlation analysis, a small negative effect size was evident between Club Velocity and VelCPyED (r = -0.25, p = 0.310). With one golfer removed (denoted as 2 in figure 5.8) there was no relationship evident (r = 0.17, p = 0.310). However, with the golfers represented by 1 (1 golfer) and 3 (2 golfers) removed, a medium non-significant effect existed (r = -0.30, p = 0.085). While the lower end of the range of VelCPyED values was negative, only two golfers produced negative values (indicated as '1' and '2' in figure 5.8). As such it might be that a VelCPyED nearer zero is the more advantageous technical trait rather than a negative VelCPyED. However this was not clear from the scatterplot with golfers with the higher Club Velocities showing a wide range of values for VelCPyED. This is evident in figure 5.8 examining the top three Club Velocities which were associated with values ranging from -0.6 to 1.8 $m.s^{-1}$.

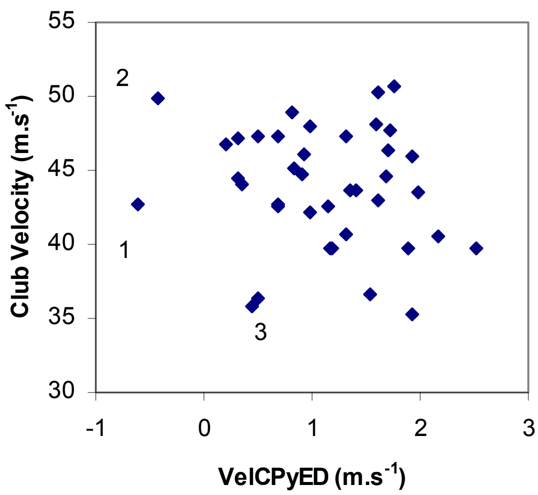


Figure 5.8: Scatterplot of VelCPyED and Club Velocity. Possible outliers denoted as 1, 2 (both one datapoint each) and 3 (two datapoints).

5.4.4.1 On the inclusion of Age in regression analysis

The association of Club Velocity and Age in regression analysis presented statistical problems in this study. While it was an important finding in itself, it was not the focus of the study. The problem arose from Age not only correlating with Club Velocity but also Handicap. As such, when Age was accounted for in correlations and regression analysis, some skill level information is likely to have been eliminated as well. The effect would have been to reduce the variance accounted for by CP parameters (i.e. they seem less influential than they were). In multiple regression analysis, the first independent variable (IV: Age in this regression) is allocated the variance accounted for that would be the case if simple correlation was performed (i.e. its 'full' amount). However, the remaining IV's are only considered in terms of the variance remaining. The other problem arising from Age correlating with performance was it precluded curvilinear analysis. This analysis was performed on an individual basis in study 3 but remains a necessary future direction for group-based analysis.

As mentioned, a number of methods of eliminating the effect of Age were attempted. Transformation was experimented with but with only $N = 39$ and $N = 19$, the choice of transformation algorithm could not be confidently determined. The best statistical and most useful practical option would have been to develop different models for different Age groups. This was performed for the under 40 year old Front Foot subset group in the next section (5.4.4.1.1). Subject numbers in the Front Foot over 40 year group and the Reverse group were too small for this analysis to be performed ($N \leq 10$ for all other subsets of Age).

5.4.4.1.1 SUBSET ANALYSIS: FRONT FOOT GOLFERS UNDER 40 YEARS OF AGE

Analysis of different Age categories indicated that 40 years of age was a key point, after which Club Velocity declined in this study. Brown *et al.* (2002; using data from this study; $N = 40$) found that no association existed between Age and Club Velocity using golfers under 40 years of age. Lockwood *et al.* (1998) also identified 40 years of age as the point of decline in skill level as indicated by an increase in handicap. Based on these findings, the 40 years of age cut-off was considered a valid partition for this study.

The same analysis used for the Front Foot group ($N = 39$) was repeated for the under 40 year old Front Foot group ($N = 29$), with the exception that bivariate rather than partial correlations were performed (as Age did not correlate with Club Velocity, $r = -0.21$, $p = 0.263$). Table 5.40 presents selected correlations between weight transfer parameters and Club Velocity.

Table 5.40: Selected correlations between Club Velocity and Handicap, Age and CP parameters for the Front Foot under 40 years old group ($N = 29$).

	Overall				Within-dataset distribution	
			Confidence Limits			
	r	p	2.5%	97.5%	$p = 0.05$	$p = 0.01$
Handicap	**-0.58**	**0.002**	-0.38	-0.74	-0.32	-0.48
Age (years)	-0.21	0.263	-0.35	0.08	-0.41	-0.52
VelCPyED ($m.s^{-1}$)	-0.11	0.555	-0.37	0.14	-0.35	-0.43
VMaxCPy ($m.s^{-1}$)	**0.61**	**<0.001**	0.43	0.77	0.33	0.41
CPyR (m)	**0.53**	**0.003**	0.31	0.71	0.36	0.42

Similar to the full Front Foot Group ($N = 39$) analysis, VMaxCPy and CPyR were significantly correlated with Club Velocity. However, VMaxCPy demonstrated a larger r-value than CPyR and as such was chosen ahead of CPyR for entry into the best regression analysis. Also different was the regression chosen from best subsets analysis included only VMaxCPy (VelCPyED was not included), and accounted for a larger amount of variance in Club Velocity than was indicated by the original analysis (R^2 = 0.38 compared with R^2 = 0.28). This might have been due to differences between the full group and subset group. It might also have been due to Age eliminating some skill factor (important information) from the analysis as discussed in the previous section.

VelCPyED was not significantly correlated with Club Velocity for the under 40 years old group (N = 29), nor was any effect size evident (r = -0.11, p = 0.555). Further, VelCPyED was not included in the best regression for the under-40 years subgroup. As well, VelCPyED was not significant when included in regressions with either CPyR or VMaxCPy (calculated for comparison purposes. CPyR and VelCPyED: Change in R^2 due to VelCPyED = 0.02, p = 0.290; VMaxCPy and VelCPyED: Change in R^2 due to VelCPyED = 0.02, p = 0.340). Clearly, this parameter was not important to the under 40 years subgroup.

Based on the findings of the under-40 years group, Age had some influence on the analysis of Club Velocity and CP parameters although the extent is not clear. While both VMaxCPy and CPyR were still significant and CPyR returned the same r-value (r = 0.53), the effect for VMaxCPy was larger in the subset analysis (r = 0.61 compared with 0.46). As well, the regression analysis was stronger and included VMaxCPy rather than CPyR (VMaxCPy was included, R^2 = 0.38 compared to 0.28). This difference might have been due to differences between the full group and subset group or due to Age eliminating some skill factor from the analysis, although it did not change the fact that both CPyR and VMaxCPy were related to Club Velocity for Front Foot golfers (rather the strength of association was altered).

5.4.4.2 Evaluation of the regression process

To evaluate the process of using cluster analysis to reduce parameter numbers and then use best subsets to choose the regression for analysis, stepwise regression was conducted for both groups. This was performed using Age in the first block and all CP parameters in the second block. This was performed to identify if a better regression had been missed due to the clustering and best subsets processes.

The Front Foot group stepwise regression included both CPyR and VelCPyED. As these results have been presented already they are not repeated here (see Results section 5.3.6.3). This confirmed that the process used in this study found the best regression in the data for the Front Foot group. While VelCPyED was not significant in replication analyses or in the under 40 years subgroup analysis, both best subsets and stepwise regression returned the same result. As such, if it was considered a limitation of the best subsets approach that a non-robust parameter was included, then the same limitation is shared by stepwise regression.

The Reverse group required more thorough analysis due to the smaller *N* and outliers. With all data analysed, the stepwise regression included CPy%ED and tMinCPy (table 5.41). However, while all regressions were significant (using the same process of eliminating cases if they failed DFIT and residual analysis), only CPy%ED was significant in the final regression (i.e. when all cases passed diagnostic tests). The analysis was repeated with outliers eliminated prior to analysis. With outliers removed, CPy%LB and VelCPyBC were included in stepwise regression analysis, as identified from the process used in this study.

Table 5.41: Results of stepwise regression for the Reverse group.

	Initial Analysis		Final solution (all cases passed diagnostic tests)	
	R^2/Change in R^2	p	R^2/Change in R^2	p
Regression	0.79	<0.001	0.91	<0.001
Age	0.62	<0.001	0.73	<0.001
CPy%ED	0.10	0.023	0.15	0.005
tMinCPy	0.07	0.040	0.03	0.191
N	19		15	
Outliers removed prior to analysis				
	R^2/Change in R^2	p	R^2/Change in R^2	p
Regression	0.83	<0.001	0.95	<0.001
Age	0.63	<0.001	0.69	<0.001
CPy%LB	0.16	0.007	0.20	<0.001
VelCPyBC	0.07	0.024	0.06	0.048
N	17		14	

The second level of decision making that was evaluated for the Reverse group was the choice of regression from best subsets. Recalling, CPy%LB and VelCPyBC were chosen as this combination produced one of the best regressions with all data and clearly the best with outliers removed. As the remaining parameter, CPy%ED, appeared in stepwise regression, it was also considered appropriate to evaluate regressions with CPy%ED included for comparison. From this analysis, the regression including CPy%LB and VelCPyBC produced the largest R^2 value, the largest change in R^2 value due to CP parameters and maintained all CP parameters as significant in the final solution. Table 5.42 summarises this analysis.

Table 5.42: Regressions for all combinations of CPy%LB, CPy%ED and VelCPyBC for the Reverse group.

	All Data included (N = 19)		Final solution (all cases passed diagnostic tests)	
	R^2/Change in R^2	p	R^2/Change in R^2	p
Full Regression	0.72	<0.001	0.95	<0.001
Age	0.64	<0.001	0.69	<0.001
CPy%LB	0.04	0.191	0.20	<0.001
VelCPyBC	0.04	0.186	0.06	0.048
N	19		14	
Full Regression	0.62	<0.001	0.91	<0.001
Age	0.62	<0.001	0.72	<0.001
CPy%LB	0.02	0.248	0.12	0.009
CPy%ED	0.10	0.060	0.07	0.011
N	19		16	
Full Regression	0.73	<0.001	0.91	<0.001
Age	0.62	<0.001	0.73	<0.001
CPy%ED	0.10	0.028	0.17	0.005
VelCPyBC	0.00	0.456	0.01	0.456
N	19		15	
Full Regression	0.75	<0.001	0.95	<0.001
Age	0.68	<0.001	0.75	<0.001
CPy%LB	0.04	0.212	0.16	0.013
CPy%ED	0.03	0.251	0.04	0.145
VelCPyBC	0.00	0.624	0.00	0.456
N	19		14	

It is difficult to evaluate the regression analysis process with the Reverse group due to influential cases affecting the data. Depending on how outliers were treated, different results were produced. However, the best regression was obtained using the process of data reduction (cluster analysis of parameters) and best subsets regression in this study. This is based on three arguments. First, stepwise regression produced the combination of Age, CPy%LB and VelCPyBC with outliers removed and all parameters were still significant after all influential cases (i.e. cases that failed diagnostics) were removed. Second, while stepwise regression produced an equation including Age, CPy%ED and tMinCPy with all data included (N = 19), only CPy%ED was significant when influential cases were removed, indicating tMinCPy was included only due to outliers. As well, from table 5.42, the R^2 value for the regression (with influential cases removed; $R^2 = 0.91$, CP parameter change in $R^2 = 0.18$) was not as large as that for the regression including Age, VelCPyBC and CPy%LB (with influential cases removed; $R^2 = 0.95$, CP parameter change in $R^2 = 0.26$). Third, no combination of CP parameters from best subsets regression (CPy%LB, CPy%ED and VelCPyBC) produced as strong a regression with all parameters significant, indicating the best subsets choice was the correct one.

5.4.5 Summary

For the Front Foot and Reverse groups, CPy% was positioned at 56-57% of the distance from the back foot to the front foot at TA. Interestingly, CPy% was moving towards the front foot rather than the back foot at TA for 71% of golfers tested. This technical trait might be related to the front leg 'press' reported by Wallace *et al*. (1990) where weight is moved to the front foot briefly at the start of backswing. There was no relationship between Club Velocity and position or velocity of CPy at TA for either group.

CPy% then moved towards the back foot through MB to LB. Position of CPy at LB was related to Club Velocity for the Reverse group but not the Front Foot group. The result indicated that for the Reverse group, a position further from the front foot at LB was related to a larger Club Velocity. Positioning CPy% at MB and rate of change of CPy at MB and LB were not related to Club Velocity for either group. At TB, CPy was positioned at 21% for the Front Foot group and 26% for the Reverse group of the distance from back to front foot and was moving towards the front foot for both groups. Positioning or rate of CPy movement at TB was not related to Club Velocity for either group. Maximum CPy velocity was similar for both groups (2.5 $m.s^{-1}$) and occurred between TB and ED for most golfers. A larger maximum CPy velocity, or a more rapid weight transfer towards the front foot, was associated with a larger Club Velocity for the Front Foot group but not the Reverse group. The time that it occurred was not correlated with Club Velocity for either group.

For both groups, CPy continued to move towards the front foot to ED. For the Reverse group, a medium positive effect existed between CPy%ED and Club Velocity but this was not significant ($p = 0.073$). For the Front Foot group, CPy velocity at ED returned a small non-significant negative effect with Club Velocity and was also included in the best subsets regression. However, this parameter was not robust in regression analysis and did not appear in the regression calculated from Front Foot under 40 years old golfers so requires more testing to substantiate.
After ED, the styles differed with the Front Foot continuing to move CPy% to the front foot (81% at BC) while the Reverse group moved weight towards the back foot (53% at BC). Position of CPy% was significantly different at MD, BC and MF ($p < 0.001$) but no relationship existed between Club Velocity and CPy% at these events for either group. However, a greater rate of CPy movement towards the back foot at ball contact was related to Club Velocity for the Reverse group. This was an unexpected finding as not only did the Reverse group position weight in a midstance position at BC rather than on the front foot, but moving CPy more rapidly towards the back foot was advantageous. This finding is in direct conflict with the coaching literature and requires that weight transfer be reassessed for coaching.

For the Front Foot group, CPy% remained predominantly on the front foot at MF (70%) compared with a position predominantly on the back foot for the Reverse group (43%). A large non-significant negative effect existed between CPy%MF and Club Velocity for the Reverse group indicating a position nearer the back foot at MF was related to a larger Club Velocity (p = 0.015). Maximum CPy (87%) occurred for the Front foot group between BC and MF and significant differences existed between groups for both the time (later in the swing) and the magnitude of maximum CPy% (larger). However, none of these parameters were significantly related to Club Velocity for either group.

The overall range of CPy movement during the swing was significantly larger for the Front Foot group compared with the Reverse group (p < 0.001). As well, a larger range of CPy movement was associated with a larger Club Velocity for the Front Foot group but not the Reverse group. However, only a small effect existed for the relationship between Club Velocity and CPy normalised to stance width. Stance width itself was also related to Club Velocity. These results indicated that adopting a wider stance to increase weight transfer in absolute terms was advantageous to Club Velocity at ball contact for the Front Foot group. Minimum CPy occurred at a similar time for both groups (near LB) but maximum CPy% occurred much later in the swing for the Front Foot group (0.01 s after BC compared with -0.23 s before BC). The Front Foot group produced a significantly smaller minimum CPy% (p = 0.020) and significantly larger maximum CPy% (p < 0.001). However, none of these parameters were correlated to Club Velocity for either group.

Regression analysis indicated that for the Front Foot group CPy range and the velocity of CPy when the club was vertical during downswing (ED) predicted 28% of the variance in Club Velocity at ball contact. However in the under 40 years subset, maximum CPy velocity was included in the regression and accounted for 38% of the variance in Club Velocity. For the Reverse group, position of CPy% at LB and CPy velocity at BC combined to predict 23% of the variance in Club Velocity and both were robust.

5.4.6 General discussion

This study has highlighted the importance of identifying if different movement strategies or styles exist within the same skill prior to performance analysis whether it be for scientific research or in coaching. Differences existed in the weight transfer parameters that were associated with Club Velocity for the two styles identified in study 1. For the Front Foot group, a larger range of CPy and a larger maximum CPy velocity were associated with a larger Club Velocity at ball contact. However, these parameters were not associated with Club Velocity for the Reverse group, for which a larger CPy%LB and more negative VelCPyBC were related to Club Velocity. In turn, these parameters were not related to Club Velocity for the Front Foot group.

If the different movement strategies had not been identified prior to correlation analysis and the group treated as one (N = 62), the effects between Club Velocity and CPy%LB (r = 0.02, p = 0.861), VelCPyBC (r = -0.08, p = 0.558) and CPyR (r = 0.17, p = 0.185) would not have been identified. Further, while the effect between Club

Velocity and VMaxCPy was significant for the full group ($N = 62$, $r = 0.38$, $p = 0.003$), this parameter was not significant for the Reverse group golfers and would be considered a type 1 error for that group. This finding has important coaching implications. For the Front Foot group, encouraging a wide stance, to produce a large range of weight transfer, and rapid forward movement may be useful coaching cues to increasing Club Velocity. However, these cues would be inappropriate for the Reverse group golfers, who employ a different technique to generate Club Velocity. For these golfers, a wider stance and large weight transfer is not important. Rather, encouraging a more rapid weight shift towards the back foot at ball contact and positioning weight further from the back foot at late backswing would be useful for increasing Club Velocity. As the importance of weight transfer to performance differed between the Front Foot and Reverse groups, it would seem that swing styles must be taken into account when working with a golfer.

Different styles within a skill or group being tested can reduce effect sizes, and hence power, in statistical analyses when all styles are assessed together. For example, the Front Foot group returned a strong association between CPyR and Club Velocity ($r = 0.53$, $p = 0.001$). However, when the group was treated as a whole ($N = 62$), no effect existed ($r = 0.17$, $p = 0.185$). As such, this relationship would have been missed in this study if styles were not identified prior to analysis (type 2 error for the Front Foot group). This highlights the potential danger of group-based analyses when different styles exist in the data.

Cluster analysis was a useful method of identifying different movement strategies within the golfers tested and proved to be an essential step prior to statistical performance analysis. For this study, the different results for the Front Foot and Reverse groups identified by cluster analysis indicated that different parameters were important for each group. Further, most of these results were not evident when the group was treated as one. The use of cluster analysis to identify styles might offer a compromise between group-based analysis with all styles analysed together (producing statistical errors) and individual-based analyses which lose the advantages of group-based analysis and are time consuming. Cluster analysis can identify strategies and group golfers into the appropriate style, and from this point statistical and/or coaching analysis performed.

5.4.6.1 Methodological issues

It was considered appropriate to discuss a number of measurement issues that were of importance to this study.

5.4.6.1.1 NORMALISATION

This study (and all previous studies) normalised CPy or Fz position to foot stance width. As outlined in the methods section, this involved calculating CPy and foot position and expressing CPy as a percentage of the distance between the feet. However, both normalised data and absolute data provided useful information in this study. For example, CPy range in absolute terms was strongly associated with Club Velocity for the Front Foot group ($r = 0.53$, $p = 0.001$) but when normalised to foot width, only a small effect existed ($r = 0.28$, $p = 0.087$). As significant correlations were found for both types of data (e.g. absolute: Front Foot CPyR – Club Velocity and relative: Reverse group CPy% - Club Velocity) the use of both should be considered in future studies.

Timing data was not normalised in this study. Normalising for time involves expressing the swing in terms of the percentage of the time between swing events such as takeaway or top of backswing and ball contact. For example, Neal (1998) used this approach in examining the time at which minimum CPy% occurred in the swing while Wallace *et al*. (1990) used swing events (similar to those used in this study). To compare absolute time as used in this study with normalised time, both were calculated and compared in terms of the correlation with Club Velocity. As indicated by Table 5.43, correlation coefficients changed only slightly and in all cases retained the same non-significant finding. Although not reported here regression analysis was also repeated with percentage time data rather than absolute data but none were included in the best regression. As such, in this study at least, absolute and relative values for time did not alter the result.

Table 5.43: Comparison of correlations between Club Velocity and timing parameters expressed in absolute and normalised terms for the Front Foot group (*N* = 39) and Reverse Group (*N* = 19)

	Normalised between	Reverse Group				Front Foot Group			
		Normalised time		Absolute time		Normalised time		Absolute time	
		r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
tVMaxCPy	TA-BC	0.34	0.182	0.39	0.117	-0.14	.401	-0.18	0.287
	TB-BC	0.30	0.241			-0.16	.325		
tMaxCPy%	TA-BC	0.33	0.194	0.34	0.182	-0.05	.778	-0.04	0.828
	TB-BC	0.39	0.120			-0.07	.672		
tMinCPy%	TA-BC	0.35	0.173	0.44	0.076	-0.14	.384	-0.12	0.470
	TB-BC	0.27	0.302			-0.18	.285		

Note: analyses were repeated with outliers removed for the Reverse group but as similar r-values existed between normalised and absolute for this data also, these have not been reported.

One of the potential problems of normalising time is the accurate identification of key swing events. As discussed in the methods section of study 1, identification of TA and TB, the two events used in the literature for normalisation (e.g. Neal, 1998), were the most difficult and least reliable swing events to identify. The collection of kinematic data from the golf club providing XYZ coordinates would have assisted in this identification. This has been used in previous research (e.g. Barrentine *et al*., 1994) and would make this measure more reliable. This was another factor that was considered in the decision to use absolute rather than relative timing for this study. The final issue of normalisation considered prior to this study was the use of swing events in preference to considering the swing in terms of percentage movement time from TA or TB to BC. This was discussed in study 1, section 4.2.5.2 but the relevant points are repeated here as they are important to this section. Swing events were chosen in this study for three reasons and two of these were specific to this study. First, swing events are easily understood by coaches and players. For example, while the top of backswing event is easily identified, the position of the body at 76% of the swing from TA to BC is not (76% was the mean normalised time that TB occurred for Front Foot golfers in this study). Second, there is growing evidence to suggest that using time-normalised data (e.g. normalised time between TA and BC) can have significant flaws because of issues of temporal dependency (e.g. Forner-Cordero *et al.*, 2006). The problem arises from the inherent assumption that there is no variability in the timing of events between TA and BC and that no rescaling occurs during the percentage conversion. Due to substantially different speeds of club movement near takeaway, the same normalised time can represent very different stages for different individuals. For example, TB occurred between 69% and 80% of the total swing time from TA to BC for different golfers. Comparison of normalised data between golfers, then, will be comparing very different stages of the swing; i.e. there is variability in the timing of events and rescaling does occur, rendering time-normalisation problematic.

5.4.6.1.2 USE OF CLUSTER ANALYSIS TO REDUCE THE NUMBER OF PARAMETERS

Cluster analysis provided a useful tool to reduce the number of parameters for regression analysis. The clustering of variables was supported by cross correlation analysis. Its use also was supported by stepwise reevaluation which confirmed that the best regression was obtained (section 5.4.4.2).While factor analysis is probably the more common method for performing this task, diagnostic tests available in SPSS (Kaiser-Meyer-Okin to test for the appropriateness of the data to be factor analysed) failed and so this method was not used.

The best subsets regression could have simply been applied to all 24 CP parameters to maximize the chance of finding the best combination of parameters that explained the variance in Club Velocity. However Tabachnick and Fidell (1996) suggested that if the number of cases is less than five for every parameter, then the results are too specific to the dataset tested and less generalisable to the population. A valid method might have been to use best subsets with all 24 parameters but only selecting from solutions that do not violate the 1:5 parameter:case ratio. However, this method still draws from the pool of all parameters (where the parameter:case ratio for the Front Foot group for example was only 1:1.6). This researcher considered that this might still produce the problem of limited generalisability described by Tabachnick and Fidell (1996) as, effectively, all the CP parameters are being used in the analysis. It could also be argued that selecting the strongest correlate from each cluster might also increase the chances of finding significance compared with planning to use a smaller number of parameters from the outset of testing. This might maximise the chance of significance in regressions where parameters have been chosen from clusters where a number of parameters correlated with Club Velocity. However, there were also clusters in which no parameter was correlated with Club Velocity and so this criterion allows for non-significant parameters to also be assessed. As well, the criteria for selection also considered previous research (e.g. choice of VelCPyED instead of VelCPyTA for the Front Foot group) so the choice was not based on statistical results only. These issues have not been discussed in the statistical literature to this researcher's knowledge. It is worth noting that, using a stepwise regression with all 24 parameters included for both the Front Foot and Reverse group data produced a similar result, indicating that the method used in this study still produced the best regression equation from the data. Recalling, stepwise regression was not used as the ratio of cases (golfers) to parameters was well below the 50:1 recommended by Tabachnick and Fidell (1996). With all parameters included, the case to parameter ratio was approximately 2:1 for the Front Foot group and 1:1 for the Reverse group and even with the use of cluster analysis to improve this ratio, the result was a ratio of approximately 5:1 for both groups.

5.5 CONCLUSIONS

Weight transfer was important for the Front Foot group and Reverse group style of swing.

Within the Front Foot group, a greater weight transfer range in absolute terms and a larger maximum CPy velocity towards the front foot in downswing were associated with a larger Club Velocity at ball contact. Stance width was related to both Club Velocity and weight transfer range indicating the increase in weight transfer is in part produced by increasing stance width. Weight transfer parameters accounted for 28% of the variance in Club Velocity. While CPy velocity at the early downswing event was also included in the best regression for the Front Foot golfers, it was not robust as it was not included in the two thirds subset or the subset of golfers who were under 40 years old. Further testing would be required to determine if this parameter is important. Encouraging an increased stance width, increased range of weight transfer and more rapid transfer of weight towards the front foot in downswing would be useful cues for Front Foot golfers to increase Club Velocity at ball contact.

Within the Reverse group, positioning CPy closer to the mid-foot position at late backswing (as opposed to closer to the back foot) and a larger CPy velocity towards the back foot at ball contact was associated with a larger Club Velocity at ball contact. The result at ball contact is in direct conflict with the coaching literature, which encourages a weight transfer that is moving towards the front foot in downswing and is positioned on the front foot at ball contact. Not only did the Reverse golfers position weight in a balanced position between the feet at ball contact but a more rapid transfer of weight towards the back foot at ball contact was associated with larger Club Velocities. These parameters accounted for 23% of the variance in Club Velocity. Encouraging a weight position further from the back foot in late backswing and a rapid weight shift towards the back foot at ball contact would be of advantage to Reverse style golfers.

The different result for each group has highlighted the importance of identifying if different movement strategies exist in the performance of a skill. If all data had been treated as one style, only maximum CPy velocity would have been indicated as important. While this conclusion would have been appropriate for the Front Foot golfers, it was not for Reverse golfers. As well, type 2 errors would have been made for both groups as the range of CPy movement would not have been identified as important for the Front Foot group and neither would CPy position at LB and CPy velocity at BC have been identified as important for the Reverse group. With no identification of movement strategies, statistical errors would have been made.

Future research needs to include kinematic as well as kinetic data to determine more technical aspects important within each swing style. In particular, this direction is essential for further exploring the underlying mechanics of the Reverse style of swing, not previously documented in coaching or scientific literature. The expansion of the performance measure to include accuracy and consistency of hit would also be appropriate. Lastly, individual analysis of these styles is essential. This is the focus of study 3.

CHAPTER 6

STUDY 3

CENTRE OF PRESSURE IN THE GOLF SWING: INDIVIDUAL-BASED ANALYSIS

6.1 AIMS

6.1.1 General

1. To examine the relationship between weight transfer and performance in the golf swing on an individual basis.

6.1.2 Specific

1. To determine if CPy parameters are related to Club Velocity on an individual basis.

2. To determine which CPy parameters are most often related to Club Velocity among individual golfers.

3. To apply a non-linear technique to individual-based centre of pressure data to determine if this technique holds useful information for individual-based analysis of the golf swing.

6.2 METHODS

6.2.1 Subjects

Five golfers from a range of skill levels were used in study 3 (table 6.1).

Table 6.1: Subject details for study 3 (N = 5)

Golfer	Handicap	Age (years)	Height (m)	Mass (kg)	Experience (years played)	Weight transfer style *
1	Professional	36	1.80	80.0	20	Reverse
2	3	29	1.94	81.4	15	Reverse
3	14	33	1.96	86.4	15	Reverse
4	5	24	1.74	80.2	15	Front Foot
5	Social	31	1.83	88.0	12	Front Foot

* Golfers were classified by comparing their CPy% means with that of the CPy% means for the Front Foot and Reverse groups (from Study 1) and choosing the nearer group using squared Euclidean distance (as used in Study 1).

The wide range of skill levels were intended to be similar to the range used in study 1 and study 2. Older golfers were not considered due to the possibility of fatigue influencing the N = 50 shot task.

Initially, N = 10 individual golfers were identified for testing for this study. This number was made up of six Front Foot golfers and three Reverse golfers (who participated in study 1 and 2) as well as one golfer not previously tested. This subject profile was used to approximate the 2:1 ratio of Front Foot: Reverse golfers indicated in Study 2 (recalling from study 1, Front Foot golfers N = 39 and Reverse golfers N = 19 – approximately 2:1). However, time delays in testing allowed for only N = 5 golfers to be tested which comprised N = 4 golfers previously tested (N = 2 Reverse golfers, $N = 2$ Front Foot golfers) as well as the previously untested golfer (identified as a Reverse golfer).

6.2.2 Task

The task was as for study 1 and study 2 with the exception that golfers completed 50 simulated drives (recalling each golfer performed only 10 simulated drives in study 1 and 2). Time between each hit was self-paced but required at least 45 s for the video file to be stored to the computer hard disk (see methods section 6.2.3).

6.2.3 Laboratory Set-up

The laboratory set-up was as for study 1 and 2 with the exception that swing events were identified from a Redlake Motionscope PCI 1000 high-speed video camera (Redlake, San Diego) operating at 250 Hz (compared with the Peak 200 Hz camera). This change was due to camera availability. Based on error assessments of both systems, the use of different cameras did not affect the analysis nor the comparison between Study 2 and 3. Measurement error in parameters used in this study is presented in section 6.2.4.1 and error calculations are presented in Appendix D.

The Redlake high-speed camera was connected to the force plate trigger system. It should be noted that this set-up did not 'gen-lock' or precisely align the timing of sampling between systems. Rather, it was used to provide ease of operation (recording was performed automatically) and reduce testing errors (i.e. ‘missing’ a trial due to starting recording too early or too late - a risk with this system as it had a short record time). It was noted that the trigger for the video system was somewhat unreliable, as the frame denoted as ball contact was often in error by one or two frames. While this did not affect the timing of sampling (i.e. the full swing was always recorded regardless of this error) it could not be used for synchronization with the force plate data. As such, identification of ball contact was performed manually. Ball contact for the video system and ball contact for the force plate system were normalized to zero and all timing measures were determined relative to this event. The time that each swing event occurred was also manually determined with the assistance of a grid placed on the screen (as for study 1) and referenced to ball contact.

6.2.4 CP Parameters

All CP parameters used in study 2 were also used in study 3 (repeated in table 6.2 with swing events in figure 6.1) with the exception of CPyR%. As CPyR and CPyR% were very strongly correlated for each golfer ($r > 0.94$) only CPyR was used.

Table 6.2: Parameters used in study 3 to assess the relationship between CP and Club Velocity in the golf swing on an individual basis

Performance	
Club Velocity	Immediately before ball contact (m.s^{-1})
Descriptive	
Handicap	
Age	Years (yr)
Height	m
Mass	kg
CP Displacement	
At each swing event	Relative to the distance between the feet (%)
CPy%TA	
CPy%MB	
CPy%LB	
CPy%TB	
CPy%ED	
CPy%MD	
CPy%BC	
CPy%MF	
CPy Velocity	
At downswing events	Instantaneous velocity (m.s^{-1})
VelCPyTA	
VelCPyMB	
VelCPyLB	
VelCPyTB	
VelCPyED	
VelCPyMD	
VelCPyBC	
VelCPyMF	
Other CP parameters	
VMaxCPy	Maximum CPy Velocity (m.s^{-1})
tVMaxCPy	Time of VMaxCPy relative to ball contact (s)
MaxCPy%	Maximum CPy% - closest position to front foot = 100%
tMaxCPy%	Time of MaxCPy relative to ball contact (s)
MinCPy%	Minimum CPy% - closest position to back foot = 0%
tMinCPy%	Time of MinCPy - relative to ball contact (s)
CPyR	CPy Range (Maximum CPy – Minimum CPy) (m)

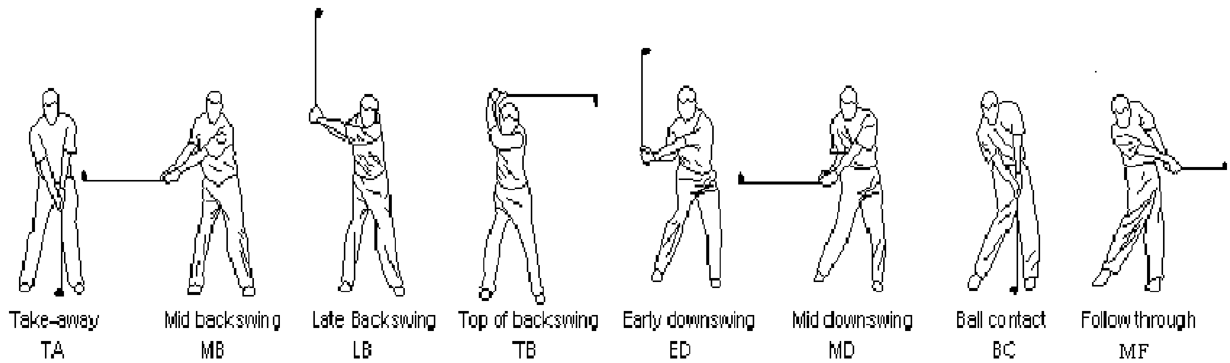


Figure 6.1: Golf swing events used in this study

6.2.4.1 Summary of error assessment

Table 6.3 presents the error associated with each parameter used in this study for a single measure and across 50 trials for each golfer. It is summarised here to allow easy reference for the reader with calculations for each parameter presented in Appendix D.

Table 6.3: Approximate error estimates for parameters used in this study (250 Hz camera). All values +/-.

Performance	Approximate Single Measure	Approximate Across 50 trials
Club Velocity ($m.s^{-1}$)	0.5	0.07
CP Displacement		
At each swing event	Relative to the distance between the feet (%)	
CPy%TA	0.6	0.08
CPy%MB	0.6	0.08
CPy%LB	0.6	0.08
CPy%TB	0.6	0.08
CPy%ED	0.8	0.11
CPy%MD	0.7	0.10
CPy%BC	0.6	0.08
CPy%MF	0.8	0.11
Average	0.6	0.08
CPy Velocity		
At each swing event	Instantaneous velocity ($m.s^{-1}$)	
VelCPyTA	0.18	0.03
VelCPyMB	0.18	0.03
VelCPyLB	0.18	0.03
VelCPyTB	0.18	0.03
VelCPyED	0.21	0.03
VelCPyMD	0.19	0.03
VelCPyBC	0.18	0.03
VelCPyMF	0.21	0.03
Average	0.19	0.03
Other CP parameters		
VMaxCPy ($m.s^{-1}$)	0.18	0.03
tVMaxCPy (s)	0.001	< 0.001
MaxCPy% (% between the	0.56	0.08
tMaxCPy% (s)	0.001	< 0.001
MinCPy% (% between the	0.56	0.08
tMinCPy% (s)	0.001	< 0.001
CPyR (m)	0.007	0.001

Average error across fifty trials for CPy% data was 0.1 % and for CPy velocity was 0.03 m.s^{-1}. Similar values were indicated for maximum and minimum parameters. As for study 2, a conservative approach to reporting data was taken with CPy% reported in units of 1% and velocity in units of 0.1 m.s^{-1} with data reported to another decimal place if required.

6.2.5 Statistical Analysis

6.2.5.1 Importance of weight transfer on an individual basis

To examine the relationship between weight transfer, as indicated by CP parameters, and performance, as indicated by Club Velocity at the instant before ball contact, on an individual basis, linear (1st order), quadratic (2nd order) and cubic (3rd order) polynomial curves were fitted to each CP parameter – Club Velocity relationship. The analysis was performed using the curve estimation option in SPSS 13.0.

To determine which curve best represented the relationship, a combination of alpha level, visual inspection of scatterplots and the statistical test presented by Hayes (1970) were used. First, R^2 and p-values were examined for each curve estimation, with an alpha level of p = 0.05 set to indicate significance. Second, scatterplots were visually inspected to see if any relationship was evident between each CP parameter and Club Velocity. Third, the significance test presented by Hayes (1970) was used to assess if the increase in R^2 between linear, quadratic and cubic fits was significant (equation 6.1). If the *F-ratio* produced by this equation was significant at $p = 0.05$ then the higher order polynomial was considered to represent a significantly different (i.e. better) fit. This method has been used in optimization studies previously examining polynomial curve fitting (Best, 1995) and was used here to add more objectivity to this decision.

$$F = \frac{\text{Square Error}_{(a-1)} - \text{Square Error}_{(a)}}{\text{Mean Square Error}_{(a)}} \qquad \text{Equation 6.1}$$

Where a = order of polynomial (linear = 1^{st}, quadratic = 2^{nd}, cubic = 3^{rd})

As mentioned, the alpha level was set at $p = 0.05$ to indicate significance. As there were a large number of analyses performed, reduction in alpha levels using Bonferroni adjustment was considered. However, while the large number of analyses will tend to capitalize on chance findings (type 2 errors) this was offset by the narrow ranges for parameters that can be expected when performing individual analysis. This will reduce effect sizes in correlation and regression type analysis (Coleman, 1999; Ball *et al.*, 2003a; Ball *et al.*, 2003b). So while the alpha level was seen as somewhat liberal, as effect sizes were likely to be reduced and as this work was largely exploratory, its use was considered reasonable.

While effect sizes have been discussed in the literature for linear relationships such as correlations and for group comparisons such as ANOVA (e.g. Cohen, 1988), to this researcher's knowledge this topic has not been discussed in relation to higher order polynomial fits. In the absence of any guidelines, the thresholds used for linear relationships have been used for quadratic and cubic fit as well (small: $R^2 \geq 0.04$, medium: $R^2 \geq 0.09$, large: $R^2 \geq 0.25$: converted from r-values presented by Cohen, 1988). For higher order polynomial fits, these thresholds are liberal as the r-value for increasingly larger orders will always be larger than the lower order.

Bivariate outliers were examined for each relationship using scatterplots. Where appropriate analyses affected by outliers were examined with and without these outliers.

The available software tools did not allow for the automation of calculation of bootstrapping for quadratic and cubic fits and so this analysis was not performed in this study.

6.3 RESULTS

Table 6.4 reports means and standard deviations of Club Velocity and CP parameters for the five individual golfers tested.

Table 6.4: Mean parameter values for individual golfers (*N* = 50 shots each).

	Reverse Group Golfers						Front Foot Golfers			
	Golfer 1 (Professional)		Golfer 2 (HCP = 3)		Golfer 3 (HCP = 14)		Golfer 4 (HCP = 5)		Golfer 5 (Social)	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Club Velocity ($m.s^{-1}$)	45.4	0.7	47.5	1.2	43.0	0.8	48.2	0.5	47.7	0.9
CP Displacement (% between the feet)										
CPv%TA	52	5	56	2	49	1	56	1	56	1
CPv%MB	17	5	24	2	33	2	32	2	48	3
CPv%LB	16	5	25	2	26	3	20	2	31	5
CPv%TB	16	5	11	2	30	4	21	2	15	3
CPv%ED	78	4	62	4	55	3	74	2	69	4
CPv%MD	79	4	61	3	60	5	90	2	80	3
CPv%BC	70	5	59	3	60	6	99	1	74	6
CPv%MF	47	8	55	4	59	8	87	3	64	8
CPv Velocity ($m.s^{-1}$)										
VelCPvTA	-0.4	0.2	-0.1	0.1	0.0	0.0	0.2	0.0	0.0	0.0
VelCPvMB	-0.3	0.1	-0.3	0.1	-0.4	0.1	-0.7	0.1	-0.3	0.1
VelCPvLB	0.1	0.1	0.2	0.2	0.0	0.2	-0.2	0.1	-0.5	0.2
VelCPvTB	0.5	0.2	0.0	0.5	0.4	0.2	0.1	0.1	0.2	0.3
VelCPvED	0.0	0.2	0.1	0.7	0.5	0.3	1.3	0.2	1.6	0.4
VelCPvMD	-0.1	0.2	-0.2	0.3	0.4	0.3	1.6	0.2	0.0	0.5
VelCPvBC	-2.2	0.4	-0.2	0.3	-0.4	0.3	0.7	0.2	-1.8	0.9
VelCPvMF	-1.1	0.8	-0.7	0.5	0.9	0.5	-0.9	0.5	0.8	0.9
Other CP parameters										
VMaxCPv ($m.s^{-1}$)	4.0	0.3	2.9	0.3	1.5	0.4	3.2	0.2	3.4	0.4
tVMaxCPv (s)	-0.17	0.004	-0.20	0.01	-0.08	0.11	-0.16	0.01	-0.15	0.01
MaxCPv% (% between the feet)	79	4	63	4	63	6	100	2	80	3
tMaxCPv% (s)	-0.07	0.03	-0.06	0.04	-0.02	0.16	0.01	0.004	-0.04	0.01
MinCPv% (% between the feet)	13	45	10	2	24	3	18	2	13	4
tMinCPv% (s)	-0.37	0.12	-0.30	0.01	-0.46	0.09	-0.43	0.07	-0.29	0.03
CPvR (m)	0.34	0.02	0.30	0.03	0.21	0.04	0.45	0.01	0.34	0.02

6.3.1 Relationship between CP parameters and Club Velocity

Table 6.5 reports results of curve estimation analysis between CPy parameters and Velocity. For consistency, R^2 values have been used for all analyses. In total, there were 43 significant linear relationships between CP parameters and Club Velocity, 37 significant quadratic relationships and 37 significant cubic relationships. After choosing the most appropriate fit for each relationship, there were 35 significant linear relationships, ten significant quadratic relationships and three significant cubic relationships.

Table 6.5: Curve estimations between CP parameters and Club Velocity for individual golfers (N = 50 shots for each golfer).

		Golfer 1 (Professional)			Golfer 2 (HCP = 3)			Golfer 3 (HCP = 14)			Golfer 4 (HCP = 5)			Golfer 5 (Social)		
		Linear	Quadratic	Cubic	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
CPy%TA	R^2	0.05	0.07	0.07	0.04	0.04	0.04	0.01	0.04	0.04	0.02	0.02	0.02	0.00	0.02	0.02
	p	0.112	0.202	0.200	0.150	0.340	0.341	0.409	0.419	0.410	0.325	0.325	0.620	0.652	0.566	0.561
CPy%MB	R^2	0.05	0.06	0.06	**0.10**	**0.17***	**0.17**	**0.08***	0.08	0.08	**0.12***	0.12	0.12	0.01	0.01	0.01
	p	0.124	0.259	0.444	**0.025**	**0.013**	**0.013**	**0.044**	0.132	0.131	**0.015**	0.052	0.052	0.538	0.802	0.799
CPy%LB	R^2	**0.08***	0.09	0.11	**0.30***	**0.31**	**0.31**	**0.11***	0.11	0.11	0.05	**0.16***	**0.16**	0.03	0.05	0.05
	p	**0.042**	0.129	0.136	**<0.001**	**<0.001**	**<0.001**	**0.021**	0.063	0.063	0.126	**0.016**	**0.016**	0.207	0.319	0.274
CPy%TB	R^2	0.03	0.03	0.04	**0.24**	**0.33***	**0.33**	**0.17***	**0.17**	**0.17**	0.01	0.02	0.02	0.01	0.05	0.06
	p	0.269	0.539	0.561	**<0.001**	**<0.001**	**<0.001**	**0.003**	**0.012**	**0.012**	0.584	0.678	0.683	0.481	0.335	0.392
CPy%ED	R^2	0.00	0.02	0.02	**0.42***	**0.43**	**0.43**	**0.39***	**0.39**	**0.39**	**0.10***	0.10	0.10	0.02	0.02	0.03
	p	0.649	0.689	0.696	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.024**	0.074	0.074	0.313	0.557	0.549
CPy%MD	R^2	0.04	0.05	0.05	**0.21***	**0.21**	**0.21**	**0.51***	**0.51**	**0.51**	0.03	0.05	0.05	0.07	0.07	0.07
	p	0.168	0.317	0.315	**0.001**	**0.003**	**0.003**	**<0.001**	**<0.001**	**<0.001**	0.195	0.288	0.290	0.068	0.188	0.188
CPy%BC	R^2	0.01	0.03	0.03	0.02	0.05	0.05	**0.40***	**0.43**	**0.43**	0.03	0.05	0.05	0.07	**0.12***	**0.12**
	p	0.437	0.535	0.528	0.401	0.275	0.278	**<0.001**	**<0.001**	**<0.001**	0.215	0.336	0.335	0.072	**0.050**	**0.050**
CPy%MF	R^2	0.01	0.01	0.01	0.04	0.05	0.05	**0.18**	**0.26***	**0.26**	0.02	0.04	0.04	**0.08**	**0.16***	**0.15**
	p	0.484	0.761	0.761	0.144	0.328	0.332	**0.002**	**0.001**	**0.001**	0.733	0.416	0.414	**0.040**	**0.018**	**0.019**
VelCPyTA	R^2	0.04	0.04	0.06	0.00	0.01	0.01	0.07	0.07	0.08	0.03	0.03	0.04	0.06	0.11	0.11
	p	0.188	0.376	0.442	0.902	0.730	0.891	0.069	0.194	0.204	0.205	0.448	0.562	0.077	0.062	0.130
VelCPyMB	R^2	0.01	0.05	0.06	0.04	0.06	0.06	0.00	0.09	0.09	0.00	0.00	0.01	0.01	0.03	0.07
	p	0.563	0.315	0.405	0.192	0.255	0.403	0.956	0.105	0.142	0.917	0.921	0.865	0.623	0.492	0.365
VelCPyLB	R^2	0.00	0.00	0.05	**0.15***	**0.19**	**0.21**	0.04	0.04	0.07	0.03	0.03	0.04	0.00	0.04	0.09
	p	0.920	0.893	0.499	**0.005**	**0.008**	**0.014**	0.161	0.373	0.372	0.223	0.467	0.646	0.707	0.430	0.222
VelCPyTB	R^2	0.08	0.10	0.10	**0.28***	**0.29**	**0.29**	**0.16***	**0.17**	**0.18**	0.02	0.02	0.02	**0.27***	**0.27**	**0.28**
	p	0.054	0.098	0.171	**<0.001**	**<0.001**	**0.001**	**0.004**	**0.011**	**0.030**	0.295	0.581	0.776	**<0.001**	**0.001**	**0.002**
VelCPyED	R^2	**0.20***	**0.20**	**0.28**	0.06	0.08	0.09	**0.30***	**0.31**	**0.31**	0.00	0.08	0.08	0.05	0.05	0.13
	p	**0.001**	**0.006**	**0.002**	0.076	0.134	0.203	**<0.001**	**<0.001**	**0.001**	0.987	0.149	0.144	0.123	0.283	0.102
VelCPyMD	R^2	0.01	0.02	0.02	**0.34***	**0.34**	**0.37**	0.08	0.11	0.12	**0.14***	**0.19**	**0.19**	0.05	0.08	**0.15***
	p	0.433	0.711	0.858	**<0.001**	**<0.001**	**<0.001**	0.052	0.062	0.125	**0.007**	**0.007**	**0.007**	0.113	0.126	**0.050**
VelCPyBC	R^2	**0.13***	**0.15**	0.15	**0.11***	0.11	0.11	0.07	0.08	0.11	0.06	0.09	0.11	0.06	0.06	0.11
	p	**0.012**	**0.029**	0.070	**0.020**	0.070	0.149	0.074	0.140	0.153	0.085	0.107	0.152	0.085	0.231	0.160
VelCPyMF	R^2	0.00	0.01	0.05	**0.12**	**0.20***	**0.20**	0.01	0.01	**0.16***	0.04	0.08	0.08	0.04	**0.14***	0.14
	p	0.665	0.823	0.480	**0.016**	**0.005**	**0.015**	0.520	0.776	**0.045**	0.168	0.149	0.284	0.180	**0.031**	0.066
VMaxCPy	R^2	0.03	0.05	0.05	**0.22***	**0.23**	**0.23**	**0.23***	**0.25**	**0.26**	0.01	0.01	0.01	0.01	0.02	0.02
	p	0.214	0.337	0.320	**0.001**	**0.002**	**0.002**	**<0.001**	**0.001**	**0.003**	0.530	0.819	0.817	0.552	0.630	0.630
tVMaxCPy	R^2	**0.08***	0.09	0.09	0.01	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.04	0.00	0.00	0.00
	p	**0.045**	0.117	0.118	0.553	0.822	0.822	0.279	0.472	0.683	0.153	0.362	0.362	0.699	0.896	0.893
MaxCPy%	R^2	0.02	0.03	0.03	**0.36***	**0.38**	**0.38**	**0.31**	**0.40***	**0.40**	0.05	0.08	0.08	0.05	0.05	0.05
	p	0.327	0.532	0.532	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.109	0.131	0.131	0.112	0.272	0.272
tMaxCPy%	R^2	**0.14***	**0.18**	**0.18**	**0.09**	**0.14**	**0.26***	**0.10***	0.10	0.10	0.01	0.02	0.03	0.02	0.05	0.12
	p	**0.008**	**0.011**	**0.013**	**0.030**	**0.029**	**0.003**	**0.026**	0.086	0.091	0.493	0.569	0.658	0.290	0.315	0.122
MinCPy%	R^2	0.05	0.05	0.06	**0.44***	**0.47**	**0.49**	**0.11***	**0.12**	**0.13**	0.02	0.06	0.06	0.07	0.08	0.09
	p	0.136	0.327	0.401	**<0.001**	**<0.001**	**<0.001**	**0.020**	**0.045**	**0.043**	0.297	0.215	0.210	0.059	0.158	0.246
tMinCPy%	R^2	0.00	0.01	0.01	0.00	0.00	0.00	0.06	0.07	0.07	0.01	0.04	0.05	**0.15***	**0.16**	**0.16**
	p	0.944	0.857	0.878	0.901	0.953	0.950	0.083	0.202	0.202	0.620	0.416	0.306	**0.005**	**0.017**	**0.016**
CPyR	R^2	**0.20***	**0.22**	**0.22**	**0.52***	**0.54**	**0.54**	**0.35**	**0.41***	**0.43**	**0.08***	**0.13**	**0.13**	**0.08***	0.08	0.08
	p	**0.001**	**0.003**	**0.003**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.050**	**0.038**	**0.037**	**0.050**	0.163	0.166

Bold type denotes significant at p < 0.05. Shaded also for clarity.

* denotes chosen relationship based on scatterplot, R^2 value and Hayes (1970) significance test

Individual golfers returned between five and 15 significant relationships between CP parameters and Club Velocity after choosing the most appropriate fit. On a general level, all golfers returned at least one significant relationship between CPy% position at swing events and Club Velocity, at least one significant relationship between CPy velocity at swing events and Club Velocity and at least one significant relationship between other CPy parameters (maximums, minimums etc) and Club Velocity. However, each golfer presented an individual-specific profile with respect to significant parameters, effect sizes, and type of relationship (linear, quadratic, cubic).

Examining the data on a parameter specific level, CPyR was related to Club Velocity for all golfers tested. For four golfers, a positive linear relationship was indicated while for Golfer 3, a quadratic best described the relationship (the linear relationship was also significant for this golfer). The next most prevalent significant association was between Club Velocity and CPy%LB (four golfers). All three Reverse golfers returned a significant linear relationship while Golfer 4 (Front Foot) returned a significant quadratic relationship. Interestingly, no relationship was significant for CPy%TA or for VelCPyTA. All remaining parameters produced significant results for at least one golfer.

Specific to swing style, for the Front Foot golfers tested, only CPyR and VelCPyMD were significantly related to Club Velocity for both golfers. For Reverse golfers, CPyR, CPy%LB and tMaxCPy% were significantly related to Club Velocity for all golfers. Also, at least a small effect was produced for VelCPyBC for all Reverse golfers.

6.4 DISCUSSION

6.4.1 Importance of weight transfer on an individual basis

Weight transfer is important on an individual basis. All individual golfers tested in this study returned multiple significant relationships at $p < 0.05$ between CPy parameters and Club Velocity. The number of significant relationships ranged from five (Golfer 4) to 16 (Golfer 3) of the 23 CP parameter – Club Velocity relationships examined. These results have provided the first scientific support for the importance of weight transfer to the golf swing on an individual basis. Further, the results provide scientific support for the emphasis on weight transfer in the coaching literature (e.g. Leadbetter, 1995; Norman, 1996).

While all golfers returned significant relationships between CPy parameters and Club Velocity, each result was individual-specific. Differences existed between golfers in terms of the number of significant results, the CP parameters that were related to Club Velocity, the strength of relationships and the nature of the relationships (e.g. positive linear, quadratic etc). For example, for the relationship between CPy%LB and Club Velocity, Golfers 1, 2 and 3 returned a significant negative linear association, Golfer 4 indicated a quadratic relationship and Golfer 5 indicated no relationship. Effect sizes encompassed large (Golfer 2: $R^2 = 0.30$), medium (Golfer 4: $R^2 = 0.16$, Golfer 3: $R^2 = 0.11$), small (Golfer 1: $R^2 = 0.08$) and practically no effect (Golfer 5: $R^2 = 0.03$). Individual-specific findings have been reported in other activities such as rifle shooting (Ball *et al.*, 2003a) and pistol shooting (Ball *et al.*, 2003b) and based on these findings, the researchers recommended intra-individual analyses for assessment of shooting activities. The finding in this study has provided strong support for a similar conclusion for the use of individual-based analysis for the examination of weight transfer in the golf swing.

CPyR was the parameter most often related to Club Velocity, with all golfers returning significant relationships. For four golfers, the relationship indicated was linear with a larger CPyR associated with a larger Club Velocity. For the fifth golfer (Golfer 3) the relationship was indicated as a quadratic. However, as indicated by figure 6.2, larger values of CPyR were still associated with larger Club Velocities and with an influential case removed (indicated by an arrow in figure 6.2) the relationship was indicated as linear (figure 6.2: $R^2 = 0.30$, $p < 0.001$). These results supported the comments of Wallace *et al.* (1990) and Koenig *et al.* (1993) that range of weight transfer was important. Importantly, this is the first statistical evidence supporting this relationship on an individual basis.

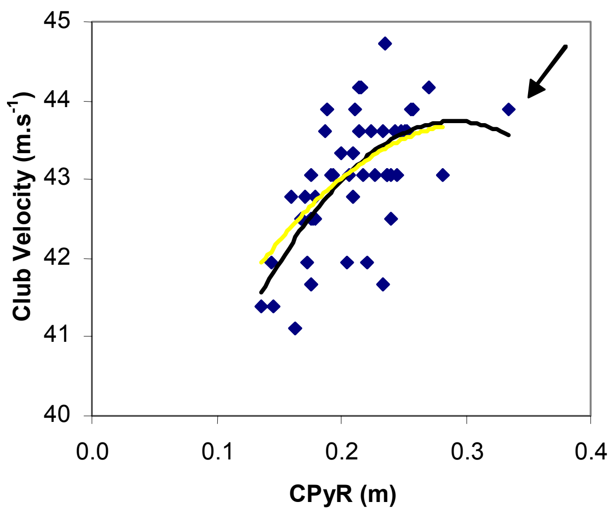


Figure 6.2: Relationship between CPyR and Club Velocity for Golfer 3 with quadratic curve (dark curve). An influential case denoted by an arrow and the lighter line indicates the quadratic with this case removed.

The mechanism underlying CPyR and Club Velocity has been discussed in study 2. Briefly repeating, it might be that increased weight transfer range assists the swing in increasing momentum developed by the whole body which is then transferred along the kinetic chain to the club head.

An important difference between individual and group-based analyses with respect to CPyR was the relationship between Stance Width and Club Velocity. For the Front Foot group, Stance Width was correlated with Club Velocity as well as CPyR (table 6.6). This prompted the suggestion in Study 2 that a cue to increase Stance Width might be a simple method of increasing CPyR and possibly increasing Club Velocity. However, Stance Width was not related to Club Velocity or CPyR for any individual golfer (Front Foot and Reverse). Clearly, increasing CPyR and not simply increasing stance width was the important element for the individual golfers tested and the cue suggested in Study 2 is more complex than was indicated by the group-based analysis. More testing examining individual golfers using different stance widths would be required to evaluate if increased stance width was a useful technical alteration on an individual basis.

Table 6.6: Relationship between Stance Width and Club Velocity and Stance Width and CPyR for the Front Foot group and individual golfers.

			Club Velocity		CPyR	
			R^2	*p*	R^2	*P*
Front Foot group		*N* = 39 golfers	**0.22**	**0.005**	**0.26**	**0.001**
Reverse Golfers						
Golfer 1	(Professional)	*N* = 50 shots	0.01	0.467	0.03	0.256
Golfer 2	(HCP = 3)	*N* = 50 shots	0.00	0.789	0.06	0.108
Golfer 3	(HCP = 14)	*N* = 50 shots	0.00	0.834	0.05	0.185
Front Foot Golfers						
Golfer 4	(HCP = 5)	*N* = 50 shots	0.00	0.806	0.06	0.102
Golfer 5	(Social)	*N* = 50 shots	0.02	0.336	0.07	0.076

Positioning of weight at specific swing events (CPy%TA to CPy%MF) is important on an individual basis. All five golfers returned at least one and up to seven significant results between Club Velocity and CPy% at the eight swing events used in this study. This indicated that different positioning of weight was associated with changes in performance. On a general level this supported the coaching emphasis on positioning of weight at key swing events such as TB and BC (e.g. Leadbetter, 1995, Norman, 1996). However, once again these relationships were individual-specific indicating different sets of factors were important for different golfers and that simply providing one 'ideal' swing coaching model is not appropriate.

CP%LB was the most consistently CPy%-at-swing-events parameter related to Club Velocity with four golfers returning a significant result. This indicated LB was an important event for weight transfer. As mentioned in study 2, this importance might be related to the proximity of LB to the onset of forces associated with downswing (decelerating the club in the backswing requires forces in the direction of the downswing). As such, positioning of weight at LB might be related to attaining the best position from which to begin downswing which in turn results in greater Club Velocity at ball contact. The time of minimum CPy% was nearer to LB than to TB for all individual golfers supporting LB being nearer the start of downswing forces. As well, Burden *et al.* (1998) noted that maximum hip and shoulder angles were achieved before the start of downswing, as defined by club movement, indicating that kinematic as well as kinetic factors point to a start of downswing forces occurring before the club begins its forwards rotation towards the ball. With four of five golfers indicating the relationship between CPy% at LB and Club Velocity was important on an individual basis (as well as on a group basis for the Reverse group in study 2), future evaluation of weight transfer in the golf swing should include late backswing as an event and/or examine the phase of the swing before the start of downswing near the LB event.

Importantly for the relationship between CPy%LB and Club Velocity, all three Reverse golfers returned a significant negative linear relationship. This result indicated that a smaller CPy%LB, or a position nearer the back foot at LB, was associated with a larger Club Velocity and that this relationship was consistent among Reverse golfers. However, while the group-based result was also significant, the direction of the relationship was positive. This difference between group-based and individual-based studies indicated that the relationship between CPy%LB and Club Velocity was more complex and would require more analysis to determine the other influential factors associated with the relationship. This difference also highlights the importance of considering individual-based analysis in evaluating weight transfer in the golf swing.

Of note in CPy% position data, for Reverse golfers all significant linear relationships from takeaway to the top of backswing were negative (CPy%TA to CPy%TB inclusive) and all in downswing events were positive (5 linear relationships; CPy%ED to CPy%BC inclusive – Golfer 1 also returned a small positive non-significant linear effect for CPy%MD). Table 6.7 details *r*-values (i.e. square root of R^2 values) with the direction of the relationship evident for all golfers. This indicated that moving CPy% nearer the back foot in backswing and nearer the front foot in downswing was associated with larger Club Velocities for individual golfers tested. Further, this suggested that a consistent pattern existed in terms of important CPy% positioning in backswing and downswing on an individual basis.

Table 6.7: Linear relationships between Club Velocity and CPy% at eight swing events for individual analysis (*N* = 50 shots for each golfer)

	Reverse Golfers					
	Golfer 1 (Professional)		Golfer 2 (HCP = 3)		Golfer 3 (HCP = 14)	
	r	*p*	*r*	*p*	*r*	*p*
CPy%TA	0.23	0.112	0.21	0.150	-0.12	0.409
CPy%MB	-0.22	0.124	**-0.32***	**0.025**	**-0.29**	**0.044**
CPy%LB	**-0.29**	**0.042**	**-0.55**	**<0.001**	**-0.33**	**0.021**
CPy%TB	-0.16	0.269	**-0.49***	**<0.001**	**-0.41**	**0.003**
CPy%ED	-0.07	0.649	**0.65**	**<0.001**	**0.62**	**<0.001**
CPy%MD	0.20	0.168	**0.46**	**0.001**	**0.71**	**<0.001**
CPy%BC	0.11	0.437	0.10	0.501	**0.64**	**<0.001**
CPy%MF	-0.10	0.484	0.21	0.144	**0.42**	**0.002**

* Linear relationship not chosen as the best fit for that set of data (both quadratic)
Bold type indicates significant at $p < 0.05$

It is possible that these results might have been simply due to golfers increasing CPyR (recalling CPyR and Club Velocity were related for all Reverse golfers). The direction of correlations (small effect or greater, $r > 0.2$) indicated that for all golfers moving CPy% nearer the back foot in backswing events and nearer the front foot in downswing events was associated with Club Velocity. This will also produce a larger CPyR. CPy% at backswing events were correlated with CPyR for two golfers (all three if CPy%TB is considered a backswing event) and CPy% at downswing events were correlated with CPyR for all golfers indicating this might be the case (table 6.8). Further, Golfer 1 and Golfer 2 returned a larger effect between CPyR and Club Velocity compared with any correlation between CPy% positioning and Club Velocity. However, Golfer 3 returned a larger effect for CPy%ED, CPy%MD and CPy%BC compared with CPyR suggesting the positioning was more important for this golfer (table 6.5 in results section 6.3.1). This means that CPy% at swing events held important information for individual golfers that could not be provided by range, maximum and minimum measures and the findings were not simply due to an increase in range of weight transfer.

Table 6.8: Correlations between CPy% in backswing and downswing events and CPyR, MinCPy% and MaxCPy%.

	CPyR			MinCPy%			MaxCPy%		
Golfer	1	2	3	**1**	2	**3**	1	2	3
CPy%TA	0.08	0.20	-0.07	**0.63**	**0.36**	0.05			
CPy%MB	**-0.32**	-0.20	**-0.45**	**0.87**	**0.49**	0.15			
CPy%LB	**-0.48**	-0.16	**-0.56**	**0.93**	**0.51**	**0.69**			
CPy%TB	**-0.32**	**-0.53**	**-0.57**	**0.94**	**0.36**	**0.86**	-0.02	**-0.64**	**-0.50**
CPy%ED	0.12	**0.61**	**0.95**				**0.94**	**0.95**	**0.62**
CPy%MD	**0.30**	**0.87**	**0.73**				**0.97**	**0.86**	**0.88**
CPy%BC	0.23	**0.88**	**0.40**				**0.91**	**0.58**	**0.92**
CPy%MF	0.05	**0.83**	0.15				**0.58**	**0.32**	**0.91**

Min and Max included for CPy%TB as it is a transition point between backswing and downswing
Bold type indicates significant at $p < 0.05$.

The rate of weight transfer at specific swing events is important on an individual basis. All individual golfers returned at least one and up to five significant relationships between Club Velocity and CPy velocity at the eight swing events used in this study. This indicated that different rates of weight transfer at specific points in the swing influenced performance. CPy velocity at specific swing events has not been reported in the scientific literature and so this represents a new and important finding. The only other study to measure rate of weight transfer performed this analysis for phases of the swing rather than at specific events but also found a significant association with Club Velocity (rate of change of Fz% between TB and FH; Robinson, 1994). Given the large number of significant results among individuals in this study as well as the only other study reporting rate of weight transfer measures finding a significant result, measurement of rate of weight transfer is essential for future testing or in applied work with golfers.

Importantly for the relationship between VelCPyBC and Club Velocity, all Reverse golfers returned a negative linear effect (although the level of significance for one golfer was only $p = 0.078$). This was similar to the group-based findings for the Reverse group and indicated that moving CPy towards the back foot more rapidly at BC was associated with a larger Club Velocity. As discussed in study 2 (section 5.4.2.2), the finding that weight is moving towards the back foot at ball contact and that a more rapid shift towards the back foot at ball contact is related to better performance represents new and important research. The only mention in the coaching literature of this occurring is in relation to the reverse pivot which is considered a technical flaw by Leadbetter (1993) and others. However, with all three Reverse golfers returning a negative linear relationship with at least a small effect, and given a similar result was found on a group-basis, this finding is strongly supported. This parameter needs to be evaluated in future work combining kinematic analysis to identify the mechanism underlying this technical trait.

While rate of weight transfer was important at swing events, neither Front Foot golfer returned a significant relationship between Club Velocity and VMaxCPy. Further, not even a small effect was evident indicating there was no relationship between these parameters for the golfers tested. This was dissimilar group-based analysis for Front Foot group ($R^2 = 0.21$, $p = 0.004$) and the under 40 years subgroup ($R^2 = 0.37$, $p < 0.001$) of which these golfers were a part. However, while no relationship existed between maximum rate of weight transfer and Club Velocity on an individual basis for Front Foot golfers, both returned a significant positive linear relationship between Club Velocity and CP velocity at swing events near where VMaxCPy occurred (Golfer 4: VelCPyMD, $R^2 = 0.14$, $p = 0.007$; Golfer 5: VelCPyTB, $R^2 = 0.27$, $p <$

0.001). As these significant results occurred at the start of or during downswing, this indicated that a rapid weight transfer was important in this phase of the swing. On a general level, this might have supported the coaching literature emphasis on a rapid weight shift towards the front foot during downswing.

Two Reverse golfers returned significant positive linear relationships between Club Velocity and VMaxCPy. For Golfer 2 and Golfer 3, a larger maximum CPy velocity, or more rapid movement towards the front foot, was associated with a larger Club Velocity. This result has provided individual-based support for the coaching literature that has reported the rapid weight shift towards the front foot in early is important (e.g. Wallace *et al.*, 1990; Koenig *et al.*, 1993; VMaxCPy occurred between TB and ED for both golfers).

No golfer (Front Foot or Reverse) returned a significant relationship between Club Velocity and CPy%TA or VelCPyTA. This indicated that neither the weight position at TA, nor the nature of the movement at TA, was important on an individual basis for the golfers tested. This was also the finding in group-based analyses in Study 2 for both Front Foot and Reverse styles, as well as all but one study in the literature (Mason *et al.*, 1995, Wallace *et al.*, 1990, but not Robinson, 1994). This finding did not support the very strong emphasis on weight position at TA in coaching texts (e.g. Leadbetter, 1995; Grant *et al.*, 1996; Norman, 1995). However as neither group nor individual golfer analyses have produced a significant result, and only one report exists in the literature of weight position at TA being important (Robinson, 1994), this conclusion was strongly supported

There are a number of possible reasons why no relationship was indicated for CPy%TA or VelCPyTA for individual golfers tested in this study. The first is that positioning weight and rate of weight transfer at TA is not important to developing Club Velocity at ball contact. A second reason, and one that is not independent of the first, is that CPy%TA and VelCPyTA are easily controllable variables or different starting positions do not affect the ensuing swing (table 6.9). A third reason might be that due to the small range of values for CPy%TA and VelCPyTA, statistical power was reduced and hence no effect was evident. On average, both produced the lowest standard deviation values of any of the swing events. As mentioned by numerous researchers (e.g. Coleman, 1999: Ball *et al.*, 2003a; 2003b) smaller ranges of values reduce effect sizes in regression analyses.

Table 6.9: Standard deviation values for individual golfers for CPy% and CPy Velocity at eight swing events (*N* = 50 shots each). Means across all individuals also included.

	Reverse Group Golfers			Front Foot Golfers		Mean
	Golfer 1 (Professional)	Golfer 2 (HCP = 3)	Golfer 3 (HCP = 14)	Golfer 4 (HCP = 5)	Golfer 5 (Social)	
Club Velocity ($m.s^{-1}$)	0.7	1.2	0.8	0.5	0.9	
CPy% at swing events (% between the feet)						
CPy%TA	5	2	1	1	1	2.0
CPy%MB	5	2	2	2	3	2.7
CPy%LB	5	2	3	2	5	3.5
CPy%TB	5	2	4	2	3	3.2
CPy%ED	4	4	3	2	4	3.4
CPy%MD	4	3	5	2	3	3.4
CPy%BC	5	3	6	1	6	4.2
CPy%MF	8	4	8	3	8	6.4
Average	*5.0*	*2.8*	*4.0*	*2.0*	*4.3*	*6.4*
CPy Velocity at swing events($m.s^{-1}$)						
VelCPyTA	0.20	0.11	0.04	0.04	0.03	2.00
VelCPyMB	0.14	0.13	0.13	0.08	0.12	2.80
VelCPyLB	0.12	0.17	0.17	0.08	0.16	3.40
VelCPyTB	0.23	0.45	0.22	0.14	0.27	3.20
VelCPyED	0.23	0.26	0.32	0.20	0.43	3.40
VelCPyMD	0.18	0.31	0.26	0.17	0.47	3.40
VelCPyBC	0.41	0.34	0.29	0.23	0.86	4.20
VelCPyMF	0.85	0.54	0.50	0.45	0.93	6.20
Average	*0.30*	*0.29*	*0.24*	*0.17*	*0.41*	*3.58*

* Extra decimal place included for velocity measures for comparison purposes.

Of note in table 6.9, the most skilled golfer tested also produced the most inconsistent weight transfer profile. The professional golfer produced the largest mean standard deviation values for CPy% parameters at swing events and the second largest mean standard deviation for CPy Velocity parameters at swing events. That the most skilled of the golfers tested in this study was the least consistent did not support Dowlan *et al.* (2001; using a subset of golfers from study 2) who reported low handicap golfers produced similar standard deviation values at TA and TB but smaller SD values at BC compared with high handicap golfers. However, the professional did produce the second most consistent Club Velocity at ball contact, indicating this golfer still produced a relatively consistent performance outcome. This combination of results was similar those reported by Arutyunyan *et al*. (1968) for elite pistol shooters who produced variable elbow and shoulder movement in aiming but maintained a high level of end point (i.e. gun barrel) control. Also of note was that only one golfer produced a smaller standard deviation at BC compared with MD. This did not support Koenig *et al*. (1993) who reported more variable Fz% values up to mid downswing and then reduced variability at BC (although for both studies, variability increased again at MF). The examination of functional variability and performance, as recommended by Davids *et al*. (2003) in motor control research, is also an important future direction for this analysis.

Swing events other than TA, TB and BC were more often related to Club Velocity in this study. As can be noted in table 6.10, overall there were a greater percentage of significant results for individuals among the swing events MB, LB, ED, MD and MF. Examining significant results specific to event (table 6.11), it can be noted that this finding was due to the lack of significance for the TA event. However one golfer returned no significant associations for TA, TB and BC (table 6.10) and so no information would have been provided to this golfer if events were limited to these. Further, for four golfers, effect sizes were larger for events other than TA, TB and BC, indicating for these golfers that other swing events were more important. This further supports the use of more swing events, and particularly events other than TA, TB and BC, in future studies examining weight transfer.

Table 6.10: Comparison of number of significant relationships between Club Velocity and the most commonly used swing events (TA, TB, BC) with other swing events used in this study

	Golfer 1	Golfer 2	Golfer 3	Golfer 4	Golfer 5	Total
TA TB BC	1	3	3	0	2	9 of 30
MB LB ED MD MF	2	7	7	4	3	23 of 50
% of total number of analyses						
TA TB BC	17%	50%	50%	0%	33%	30%
MB LB ED MD MF	20%	70%	70%	40%	30%	46%

Table 6.11: Comparison of number of significant relationships between Club Velocity and individual swing events

	CPy%	CPy Velocity	Total
TA	0	0	0
MB	3	0	3
LB	4	1	5
TB	2	3	5
ED	3	2	5
MD	2	3	5
BC	2	2	4
MF	2	3	5

The importance of positioning of weight compared to rate of weight transfer differed between backswing and downswing events. Positioning weight was more important at backswing events (MB and LB) compared to CPy velocity for the golfers tested. Conversely, rate of weight transfer was more important in downswing events (ED and

MD) compared to positioning. For positioning in the backswing, six significant relationships between Club Velocity and either CPy%MB (two golfers) or CPy%LB (four golfers) compared to only one significant relationship for CPy velocity. For rate of weight transfer in the downswing, all five golfers returned a significant relationship between either VelCPyED or VelCPyMD and Club Velocity compared with only three of five golfers returning significant relationships for CPy% positioning at the same events. Further, while the number of significant results were the same (five for CPy% positioning and five for CPy velocity at ED and MD), there were four non-significant small effects for selected relationships between CPy velocity and Club Velocity, compared with only one for CPy%. This might be suggesting a differing importance of weight transfer in the golf swing with attaining the correct position in the backswing and then achieving the correct rate of weight transfer in the downswing as factors that combine to facilitate better performance. However, this can only be seen as speculative with the limited number of individual golfers tested. This area should be examined in future individual-based analyses.

Timing of weight transfer was important for three golfers (after the relationships with measurement issues were recalculated – see section 6.4.2.1). However there was no consistency among the significant results. Only one golfer produced an effect for the relationship between tVMaxCPy and Club Velocity (Golfer 1). As well, only one golfer produced a significant result for the relationship between tMinCPy% and Club Velocity (Golfer 5: positive linear relationship) and although a small effect was produced by another golfer, this effect was not significant at $p < 0.05$ and indicated a different direction of relationship to that of Golfer 5 (Golfer 3: negative linear). Further, while two golfers produced a significant result for tMaxCPy% and Club Velocity, the type of relationship differed. For Golfer 2, a cubic relationship existed although with three cases removed, the relationship was negative linear. For Golfer 1, measurement issues indicated that regression statistics may not be appropriate for tMaxCPy% (refer to section 6.4.2.1) but swings where maximum CPy occurred later in the swing were related to larger Club Velocities, indicating different relationship to that produced by Golfer 2 regardless. As such while timing was important for three golfers, it was individual-specific.

6.4.1.1 TIMING MEASUREMENT ISSUE

The same measurement issue discussed in section 5.4.3 in study 2 existed for two individual golfers in study 3. Figures 6.3 and 6.4 show scatterplots of the problem relationships for Golfer 1 and Golfer 3 with example CPy curves of specific trials selected from each of the clusters to highlight the cause of the measurement issue.

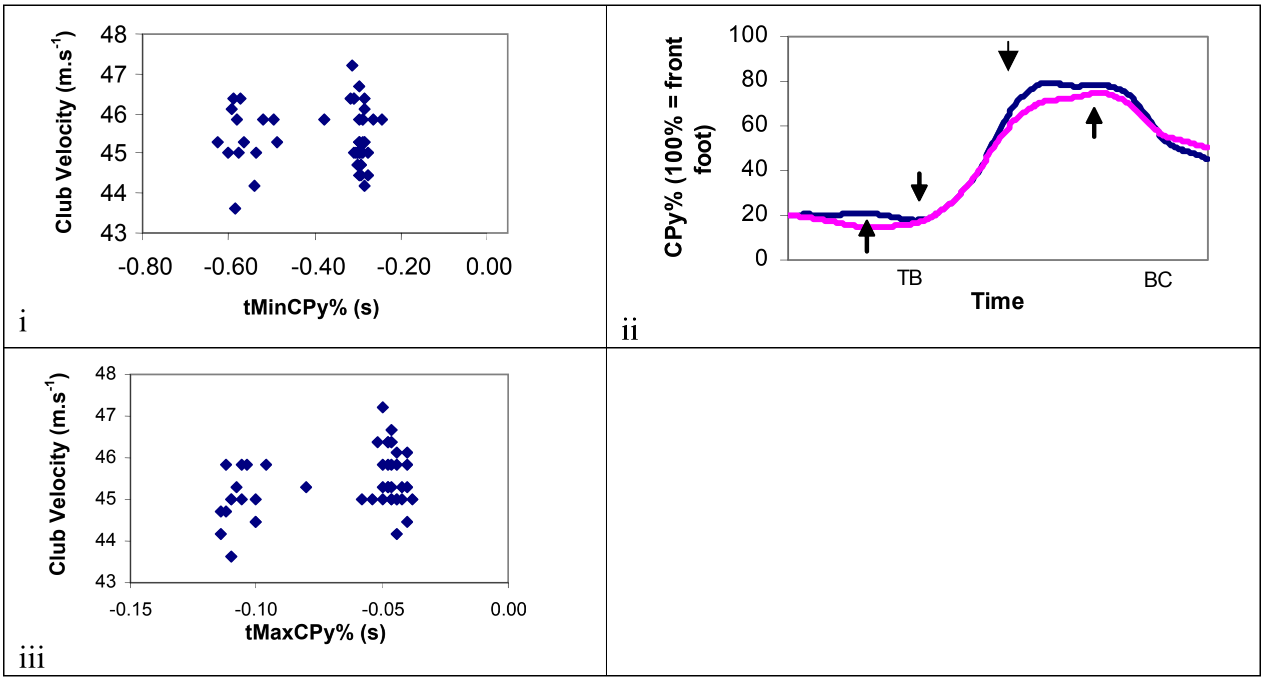


Figure 6.3: Scatterplots and example CP curves for Golfer 1 highlighting the clusters and the cause of the measurement issue for tMaxCPy% and tMinCPy%. Arrows denote sites of maximum and minimum values relating to the clusters in scatterplots.

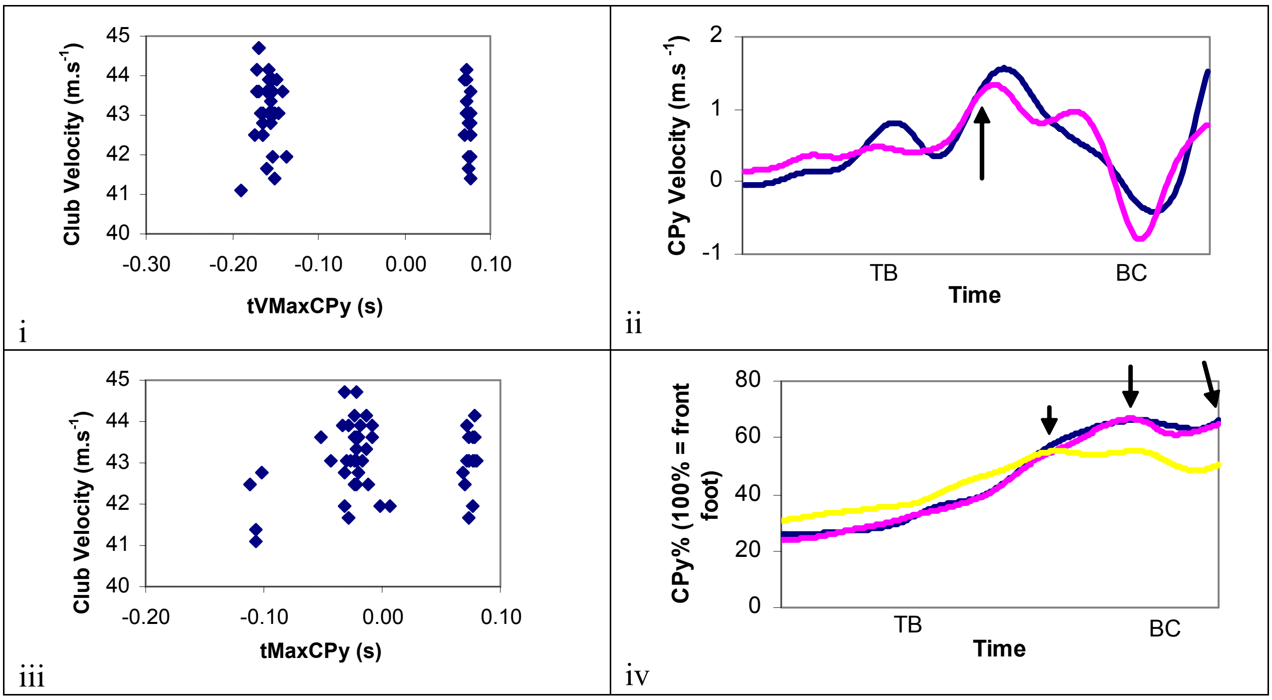


Figure 6.4: Scatterplots and example CP curves for Golfer 1 highlighting the clusters and the cause of the measurement issue for tVMaxCPy and tMinCPy%. Arrows denote sites of maximum and minimum values relating to the clusters in scatterplots.

For Golfer 1, two peaks existed across the crest of the CPy trace between TB and BC and slight differences in the shape of the curve meant that tMaxCPy% related to the first or the second peak. Also, slightly different patterning just before TB produced different minimum CPy times. For Golfer 3, both maximum CPy% and maximum CPy velocity occurred either midway between TB and BC in downswing or immediately after BC but the difference in time of maximum was associated with different 'hills' in the curve. A small third group also existed for maximum CPy and was produced in swings that showed a different patterning from the other swings (yellow line figure 6.4 iv).

Three measurement issues exist due to the clusters in figures 6.3 and 6.4. First the mean value presented in table 6.4 (Results section 6.3) did not relate to any of the peaks indicated in figures 6.3 and 6.4. Second, the clusters represent very different parts of the swing. Both these factors were an issue on a group basis as well. As well, a third issue existed for individual-based analysis relating to the use of regression statistics where clusters are evident in the data. This can affect the analysis by producing or masking effects simply due to the existence of groups of data. A similar possibility was raised in relation to clusters due to different swing styles (refer Literature Review figure 2.9 in section 2.1.5.4).

To evaluate these factors, within-cluster and between-cluster analyses were performed. Data were grouped into respective clusters (i.e. the clusters evident in scatterplots in figures 6.3 and 6.4) and the relationship between Club Velocity and the CPy timing parameter was then examined within each of these clusters. Other parameters that might have been influenced by this issue were also re-examined (i.e.

MaxCPy%, MinCPy%, CPyR and VMaxCPy). Also for Golfer 3, maxima were re-evaluated between TA and BC (rather than MF), which eliminated the clusters produced after BC.

For Golfer 1, clusters in tMaxCPy affected the analysis. Neither within-cluster relationship between Club Velocity and tMaxCPy was significant indicating the original significant result was influenced by the existence of the clusters in the data (table 6.12). However, Club Velocity was significantly different between clusters indicating that swings that produced a later maximum for CPy% also produced the larger Club Velocities on average. As such, while a statistical issue might have made the use of regression statistics questionable, the information produced (i.e. that a later maximum CPy% was advantageous for Golfer 1) remained. For this relationship the more appropriate method of analysis would be to compare clusters rather than examine regression statistics.

Table 6.12: Within-cluster and between-cluster analysis for CP parameters affected by tMaxCPy% for Golfer 1

			tMaxCPy% (s)	MaxCPy% (%)	CPyR (m)	Club Velocity (m.s^{-1})
Cluster 1	$N = 16$	Mean	-0.11	78	0.33	45.0
Cluster 2	$N = 34$	Mean	-0.05	80	0.35	45.6
ANOVA		F	1021	2.7	4.5	7.8
		p	<0.001	0.105	0.038	0.008
Relationship within Cluster 1		R^2	0.06	0.01	0.13	
		p	0.351	0.719	0.173	
Relationship within Cluster 2		R^2	0.03	0.01	0.15	
		p	0.341	0.504	0.024	

Note: all chosen relationships were linear

The original results for the relationships between Club Velocity and MaxCPy% and between Club Velocity and CPyR were supported for Golfer 1. Neither cluster returned a significant within-cluster relationship between MaxCPy% and Club

Velocity, supporting the original non-significant analysis. For CPyR, both within-cluster relationships with Club Velocity were positive linear, supporting the original analysis. Although the relationship for cluster 1 was not significant, the medium effect was only 0.2 less than cluster 2 with the non-significant result largely due to lower *N*. As well, a significant difference existed between clusters 1 and 2, with the larger CPyR group (cluster 2) returning a significantly larger Club Velocity.

The clusters evident in the tMinCPy% data did not influence the analysis for Golfer 1 (table 6.13). There was no significant relationship between tMinCPy and Club Velocity for either within-cluster group, supporting the non-significant original analysis ($R^2 = 0.05$, $p = 0.134$). The results for the relationships between Club velocity and both MinCPy% and CPyR were also unaffected. There was no significant relationship within either cluster for MinCPy% supporting the original non-significant analysis ($R^2 = 0.00$, $p = 0.944$). For CPyR, both within-cluster relationships were positive linear supporting the original positive linear relationship ($R^2 = 0.20$, $p = 0.001$). While the level of significance for this effect was only $p = 0.115$ for cluster 1, the R^2 value was only 0.1 below the original analysis and the non-significant finding was largely due to the lower *N*.

Table 6.13: Within-cluster and between-cluster analysis for CP parameters affected by tMinCPy% for Golfer 1

			tMinCPy% (s)	MinCPy% (%)	CPyR (m)	Club Velocity ($m.s^{-1}$)
Cluster 1	$N = 14$	Mean	-0.56	15	0.34	45.4
Cluster 2	$N = 36$	Mean	-0.30	12	0.34	45.4
ANOVA		F	963	3.2	0.1	0.0
		p	<0.001	0.082	0.725	0.839
Relationship Cluster 1		R^2	0.00	0.07	0.19	
		p	0.954	0.354	0.115	
Relationship Cluster 2		R^2	0.04	0.04	0.22	
		p	0.283	0.245	0.004	

Note: all chosen relationships were linear

The clusters evident in tVMaxCPy data did not influence the analysis for Golfer 3. As can be noted in table 6.14, there were no significant relationships between tVMaxCPy and Club Velocity for either cluster, supporting the original non-significant finding (R^2 = 0.02, p = 0.279). Further, the original result for VMaxCPy (positive linear, R^2 = 0.23, p < 0.001) was supported by within-cluster relationships with Club Velocity, with both clusters producing high medium effects (both positive linear) as well as when VMaxCPy was re-evaluated between TA and BC (R^2 = 0.23, p < 0.001). While the effect was not significant for cluster 2, the R^2 value was only 0.04 below the original analysis and the non-significant finding was due largely to smaller N.

Table 6.14: Within-cluster and between-cluster analysis for CP parameters affected by tVMaxCPy for Golfer 3

			tVMaxCPy (s)	VMaxCPy (m.s-1)	Club Velocity (m.s-1)
Cluster 1	N = 32	Mean	-0.16	1.5	43.1
Cluster 2	N = 18	Mean	0.07	1.6	42.9
ANOVA		F	8437	0.4	1.2
		p	<0.001	0.542	0.282
Relationship Cluster 1		R^2	0.00	0.31	
		p	0.908	0.001	
Relationship Cluster 2		R^2	0.12	0.15	
		p	0.160	0.098	
Relationship using the interval from TA - BC		R^2	0.01	0.23	
		p	0.552	<0.001	

Note: all chosen relationships were linear

Clusters affected the result for tMaxCPy% for Golfer 3. The original analysis indicated a significant polynomial relationship. However, within-cluster analysis indicated no relationship between tMaxCPy% and Club Velocity (table 6.15). Further, the relationship between tMaxCPy% and Club Velocity was not significant when MaxCPy% was evaluated between TA to BC, which eliminated cluster 1 from the analysis, and the small effect was due only to four outliers (represented by cluster 3,

$R^2 = 0.08$, $p = 0.059$, with four outliers represented by cluster 3 removed $R^2 = 0.03$, $p = 0.541$). This indicated that while it was better for Golfer 3 to produce swings where MaxCPy% occurred in cluster 2 compared to cluster 3, no relationship existed between tMaxCPy% and Club Velocity for the majority (i.e. $N = 46$) of swings.

Table 6.15: Within-cluster and between-cluster analysis for CP parameters affected by tMaxCPy for Golfer 3

			tMaxCPy% (s)	MaxCPy% (%)	CPyR (m)	Club Velocity ($m.s^{-1}$)
Cluster 1	$N = 15$	Mean	0.07	63	0.22	40.5
Cluster 2	$N = 31$	Mean	-0.02	60	0.21	41.9
Cluster 3	$N = 4$	Mean	-0.08	44	0.13	33.7
ANOVA		F	9.0	17.1	0.33	4.7
		p	0.004	<0.001	0.718	0.013
Relationship Cluster 1		R^2	0.06	0.21	0.24	
		p	0.378	0.079	0.049	
Relationship Cluster 2		R^2	0.03	0.55	0.45	
		p	0.548	<0.001	<0.001	
Relationship using the interval from TA - BC		R^2	0.08	0.52	0.41	
		p	0.059	<0.001	<0.001	

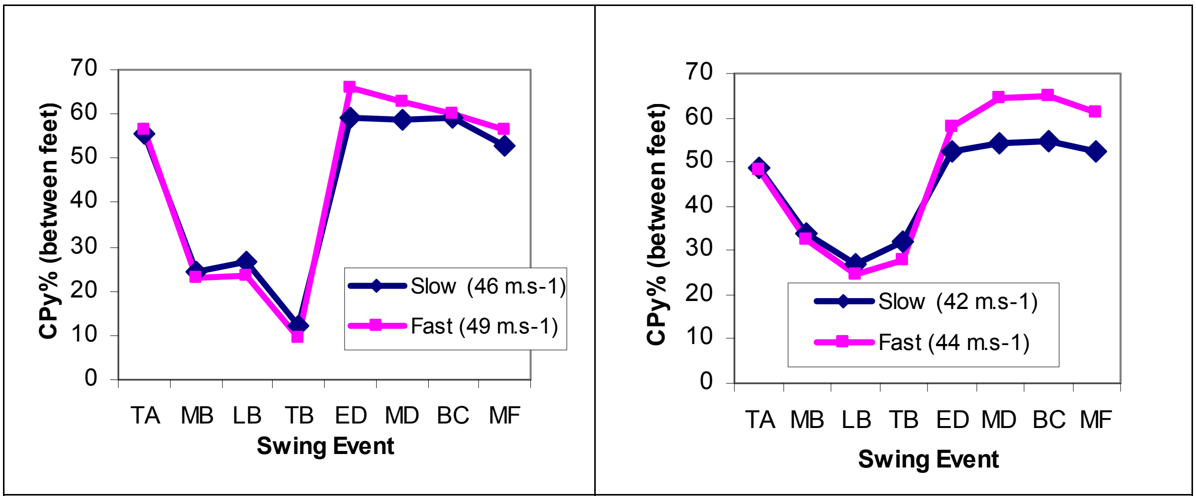
Note: all chosen relationships were linear

This measurement issue for Golfer 3 did not alter the decision that both MaxCPy% and CPyR were significantly related to Club Velocity, although the nature of the relationship indicated was different. The relationship was indicated as linear for each cluster rather than quadratic as was indicated in the original analysis (level of significance was $p = 0.079$ only for Cluster 1 for CPyR but this was due in part to low N with a medium effect evident)., However when MaxCPy% was re-evaluated for the interval from TA to BC (i.e. effectively eliminating cluster 1) the relationship with Club Velocity was a quadratic for both and returned the same R^2 values as the original analysis indicating minimal influence of the timing measurement issue on these parameters (MaxCPy%: $R^2 = 0.52$, $p < 0.001$; CPyR, $R^2 = 0.41$, $p < 0.001$). It should be noted that on a general level, all analyses indicated larger MaxCPy% and CPyR values were associated with larger Club Velocities.

On a general level, there are a number of options when treating these measurement issues. For Golfer 1 and maximum CPy%, a valid method of analysis would be to compare the two clusters. Also, examining only the interval between TA and BC was appropriate for Golfer 3 and will be particularly important where the maximum being detected is on a different 'hill'. While the differences between the swings might be of importance and should be examined, the data should not be evaluated together as substantial violations of statistical assumptions will exist (e.g. non-normal data). As well, very different aspects of the swing were measured between each data cluster for Golfer 3 for both tMaxCPy% and tVMaxCPy. It is also important to evaluate measures other than the time parameters themselves such as maximum and minimum values as these parameters might also produce different results if values from the same 'hill' are used rather than from different hills.

6.4.2 On clustering of individual golfers

The strong positive relationships between Club Velocity and both CPy%MD and CPy%BC were unusual for Golfer 2 and 3. For both, higher Club Velocities were produced in swings in which CPy% was positioned nearer the front foot at swing events that are near or at ball contact, a trait evident in Front Foot but not Reverse golfers. Figure 6.5 shows the mean of the 10 swings with the largest Club Velocities and 10 lowest Club Velocities. For Golfer 3, the differences between CPy% at TB, ED, MD, BC and MF were all significant at $p < 0.05$ while for Golfer 2, significant differences existed at LB, TB, ED and MD (table 6.16).



Golfer 2	Golfer 3

Figure 6.5: Comparison of CPy% at eight swing events for the 10 fastest and 10 slowest Club Velocities for Golfer 2 and 3.

Table 6.16: One way ANOVA comparing shots $N = 10$ largest and $N = 10$ smallest Club Velocities for Reverse Golfer 2 and Reverse Golfer 3

	Golfer 2 (HCP = 3)						Golfer 3 (HCP = 14)					
	Large Club Velocity ($N = 10$ shots)		Small Club Velocity ($N = 10$ shots)		ANOVA		Large Club Velocity ($N = 10$ shots)		Small Club Velocity ($N = 10$ shots)		ANOVA	
	Mean	SD	Mean	SD	*F*	*p*	Mean	SD	Mean	SD	*F*	*p*
Club Velocity ($m.s^{-1}$)	49	1	46	1	179.6	<0.001	44	0	42	0	214.7	<0.001
CPy%TA	56	2	56	1	1.1	0.311	48	1	49	2	0.3	0.620
CPy%MB	23	3	24	1	1.9	0.182	32	2	34	2	2.1	0.165
CPy%LB	23	3	27	2	12.3	0.003	25	3	27	2	3.1	0.098
CPy%TB	9	2	12	2	11.8	0.003	28	3	32	3	10.3	0.005
CPy%ED	66	3	59	2	32.1	<0.001	58	3	53	3	19.4	<0.001
CPy%MD	62	4	59	2	6.6	0.020	64	3	54	5	25.6	<0.001
CPy%BC	59	4	59	2	0.0	0.918	65	4	55	6	16.1	0.001
CPy%MF	54	5	53	3	0.7	0.430	61	7	52	9	6.2	0.023

Evident in figure 6.5, for Golfer 2, the slower swings were associated with CPy% pattern that maintained a similar CPy% position for ED, MD and BC. This compared with the faster swings which exhibited a CPy% position nearer to the front foot at ED and MD and a more noticeable reverse pattern. This indicated that the swings in which a reverse technique was produced were the swings that achieved the largest Club Velocity, and if the reverse pattern was not produced then slower Club Velocities resulted. This supported the allocation of Golfer 2 to the Reverse group and that the reverse patterning was important to this golfer.

Conversely, Golfer 3 did not produce a large reverse pattern for the lower or higher Club Velocity swings. Rather, the swing pattern resembled a Front Foot rather than Reverse style for Golfer 3, particularly for the faster swings. CPy% was positioned closer to the front foot at MD compared with ED while for the Reverse group mean, these values were very similar. As well, the time of maximum CPy% occurred later in the swing for Golfer 3 (94%) compared with the Reverse group in study 3 (82%). This unusual pattern for Golfer 3 prompted a post-hoc analysis of classification procedures. Reanalysis of classification using the Squared Euclidean and Pearson's correlation methods indicated that while Golfers 1, 2, 4 and 5 classified in the same group (table 6.17), Golfer 3 exhibited a variable classification. Using the squared Euclidean distance measure, which uses the absolute values of each parameter, Golfer 3 was classified as a Reverse golfer due to CPy% remaining nearer a midfoot position in both backswing and downswing (figure 6.6). However, when using the Pearson's correlation method, which examines the overall pattern of the weight transfer, Golfer 3 was indicated as a Front Foot golfer.

Table 6.17: Clustering results using the Pearson's Correlation and Squared Euclidean Distance measures for individual golfers. Bold values indicate most appropriate cluster for each golfer for each method.

	Pearson's Correlation (largest value indicates best cluster fit)		Squared Euclidean Distance (smallest value indicates best cluster fit)	
	Reverse	Front Foot	Reverse	Front Foot
Golfer 1	**0.97**	0.88	**1220**	1736
Golfer 2	**0.94**	0.88	**420**	2075
Golfer 3	0.83	**0.95**	**649**	1976
Golfer 4	0.81	**0.99**	5358	**566**
Golfer 5	0.85	**0.89**	2013	**1297**

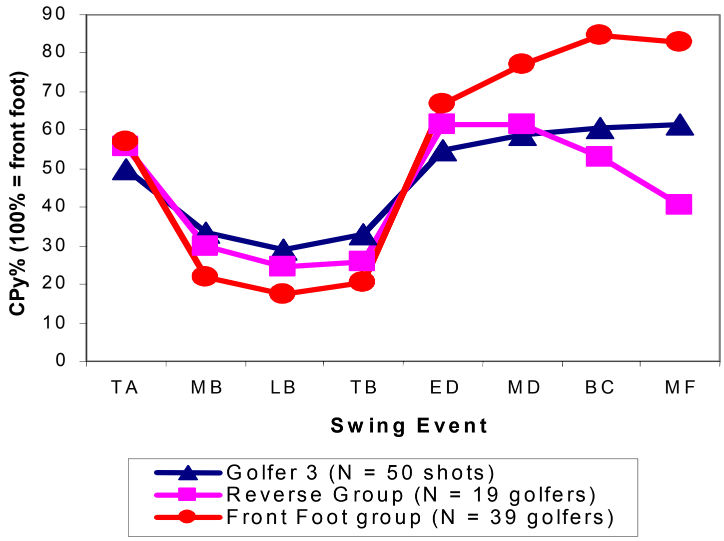


Figure 6.6: Comparison of CPy% at eight swing events for Golfer 3, Front Foot group and Reverse group.

In practical terms, it would probably be better to classify this golfer as a Front Foot style golfer when examining Club Velocity at ball contact. Swings with larger Club Velocity were nearer the Front Foot style indicating a more front foot style was better in terms of Club Velocity for Golfer 3. As well, the relationships between Club Velocity and CPy% at swing events ED, MD, BC and MF were linear and positive indicating that further increases in CPy% at these events might be of advantage (i.e. position CPy% nearer the front foot during downswing and follow through events). This technical change would further liken the pattern to that of the Front Foot group.

Four individual golfers (this study) and 92% of the major group golfers from Study 1 (section 4.4.5.3) were classified into the same cluster regardless of measure used. This would indicate that the use of cluster analysis is an appropriate method of classification for the majority of golfers. However, two golfers in Study 1 (8%) and one golfer in Study 3 (not previously tested) exhibited variable classification. As such it would seem that there will be some golfers that are difficult to classify using this method. For these golfers, a more detailed assessment of which style is most appropriate is necessary.

Possible solutions to the problem of clustering for golfers who exhibit aspects of both styles could be provided by Fuzzy clustering (e.g. Chau, 2001) or the use of more than one cluster measure. Fuzzy clustering will produce a percentage membership for each golfer for the Front Foot and Reverse styles and those golfers who indicate a high percentage clustering with more than one cluster can be indicated as requiring further analysis. However, fuzzy clustering might not have identified Golfer 3 as a potential problem if only the squared Euclidean distance was used as, observing figure 6.6, CPy% values were close to the Reverse mean for all events (and so this golfer would most likely have had a large percentage belonging value for the Reverse group). For this golfer, the use of both squared Euclidean and Pearson's correlation methods would have been needed to identify a possible classification problem. Using this method, any golfer who changes cluster with the different methods would need to be further analysed. Once golfers with classification issues are identified, the process used in this section (fastest and slowest 10 swings compared) could be implemented to provide more information on which classification would be optimal. Important information was gathered for both Golfer 2 and Golfer 3 using this process and this information supported the inclusion of Golfer 2 in the Reverse group.

6.4.3 Practical significance of results

To provide an indication of the practical significance of the findings in this study, total drive distance was approximated. Total drive distance is the distance between ball strike and where the ball finally comes to rest. Quintavala (2006) in a technical report for the US golf association reported approximate distance data predicted from Club Velocity at ball contact. This was based on experimental data obtained using a mechanical golf swing system which hit five different ball types at four (reliable) Club Velocities ranging from 40 $m.s^{-1}$ to 54 $m.s^{-1}$. These values encompassed the measured values in this study (refer table 6.18). Although the conversion factor was not reported, it could be calculated from data presented in the paper. For the purposes of this assessment, the average conversion factor for all conditions was used (five balls at four Club Velocities = 5.2 m of total drive distance for every 1 $m.s^{-1}$ of Club Velocity at ball contact. N.B. the relationship is quadratic, not linear, but in this range of club velocities a linear relationship is adequate). Results of this approximation are reported in table 6.18.

Table 6.18: Range of Club Velocity values for individual golfers. Approximate Total Drive Distance ranges also included.

	Reverse Group Golfers			Front Foot Golfers	
	Golfer 1 (Professional)	Golfer 2 (HCP = 3)	Golfer 3 (HCP = 14)	Golfer 4 (HCP = 5)	Golfer 5 (Social)
Range ($m.s^{-1}$)	43.6 – 47.2	45.0 – 50.6	41.1 – 44.7	47.5 – 49.2	45.0 – 49.7
Approximate Total Drive Distance range (min - max, m)	227 - 245	234 - 263	214 - 232	247 - 256	234 – 258
Approximate Total Drive Distance range (difference, m)	19	29	19	9	24

Golfers in this study produced approximate Total Drive Distance ranges of between 9 m and 29 m. These differences in total drive distance are substantial, particularly for the four golfers who ranged in distance of 19 m or more. Numerous golf coaching websites indicate that the difference between consecutive clubs (i.e. 3-iron to 4-iron, 4-iron to 5-iron) is between nine and 14 metres (e.g. Kelley, 2006; Meinen, 2006; Leaderboard, 2006). These ranges are similar to experimental results from the Victorian state golf team (average difference in distance between consecutive clubs = 10m, $N = 8$, Cooney, 2005, personal communication, 10 October). Using 10 m as one club length, a drive that takes the ball 20 m further (two club lengths) will require a shorter club for the next shot on par 4 holes and possible par 5 holes (e.g. a 7-iron rather than a 5iron). Given four golfers ranged 19 m or more for total drive distance from Club Velocities produced in testing, factors that contributed to this change are worth assessing.

So what does this mean for the importance of weight transfer to the golf swing? To gain some insight into the practical significance of how weight transfer contributed to the differences in Club Velocity, a post-hoc analysis of CPyR was performed (CPyR was chosen as it was significant for all golfers). For each golfer, the ten largest (termed LargeCPyR) and ten smallest (termed SmallCPyR) CPyR swings were grouped together. Mean Club Velocity and estimated ball carry values were calculated for each group and compared using a one way ANOVA. The results of this analysis are presented in table 6.19.

Table 6.19: Differences in Club Velocity and estimated ball carry between ten largest and ten smallest CPyR values for individual golfers

	Reverse Golfers						Front Foot Golfers			
	Golfer 1 (Professional)		Golfer 2 (HCP = 3)		Golfer 3 (HCP = 14)		Golfer 4 (HCP = 5)		Golfer 5 (Social)	
	Club Velocity ($m.s^{-1}$)	CPyR	Club Velocity ($m.s^{-1}$)	CPyR	Club Velocity ($m.s^{-1}$)	CPyR	Club Velocity ($m.s^{-1}$)	CPyR	Club Velocity ($m.s^{-1}$)	CPyR
LargeCPyR (mean, N=10)	46.4	0.35	49.3	0.33	44.1	0.24	49.1	0.46	49.0	0.35
SmallCPyR (mean, N=10)	44.5	0.32	45.9	0.27	41.8	0.18	47.8	0.45	46.6	0.33
Difference between LargeCPyR and Small CPyR swings	1.9	0.03	3.4	0.06	2.4	0.06	1.3	0.01	2.4	0.02
Group Difference for Total Drive Distance	10 m		18 m		12 m		7 m		12 m	
F	139.6	28.6	57.4	49.99	214.7	15.5	520.0	11.7	123.1	4.6
p	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.001**	**<0.001**	**0.005**	**<0.001**	**0.048**
Effect size (η^2)	0.886	0.614	0.909	0.735	0.923	0.463	0.967	0.394	0.872	0.204
Effect Scale	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large

All golfers returned significant differences between LargeCPyR and SmallCPyR for both CPyR and Club Velocity ($p < 0.05$). For four of five golfers, the difference in mean Total Drive Distance was larger than 10 m with the LargeCPyR group producing greater distance. As such, for these golfers, the difference in performance related to one club length. This indicated a high level of practical significance for the relationship between CPyR and Club Velocity for these four golfers.

Golfer 4 produced a mean difference in Total Drive Distance of only 7 m between LargeCPyR and SmallCPyR and a maximum difference of only 9 m (calculated from the single maximum and single minimum Club Velocity produced by Golfer 4; table 6.19). As such, this golfer did not produce a large range of Club Velocities in the N = 50 swings and so there was no possibility of finding an increase of 10 m or more for this golfer. However, a 7 m increase in distance is likely to be advantageous as the golfer is closer to the green for the second shot potentially allowing for a shorter club to be taken or bringing the green into range.

6.4.4 Use of group-based and individual-based analyses

Individual-based analysis provided information that was not evident on a group-basis in this study. Numerous parameters were significantly related to Club Velocity on an individual basis but not on a group basis. For example, both Front Foot golfers returned significant relationships between CPy% at swing events and Club Velocity while there were no significant effects on a group basis. For Reverse golfers, the relationship between CPyR and Club Velocity was related on an individual basis for all golfers while not being significant on a group basis. Using only group-based analysis, these factors would not have been identified and so not offered as possible technical aspects to alter for improved performance for these individual golfers.

Group-based analysis provided information that was not evident on an individual-basis in this study. For the Reverse golfers, a significant negative relationship was indicated between VelCPyBC and Club Velocity on a group basis but not an individual basis for Golfer 3. However, the range of values for Golfer 3 was small, reducing statistical power for the analysis. Further, the mean value for Golfer 3 for VelCPyBC was one of the least negative values produced by any golfer (3rd largest, figure 6.7). Given this combination, it might be useful for Golfer 3 to increase the negative VelCPyBC values further. Using only individual-based data, this possibility would not have been detected. This combination of results supports the recommendation of Ball *et al.* (2003a, 2003b) that both group-based and individual-based analysis is required to extract all the available information form an analysis.

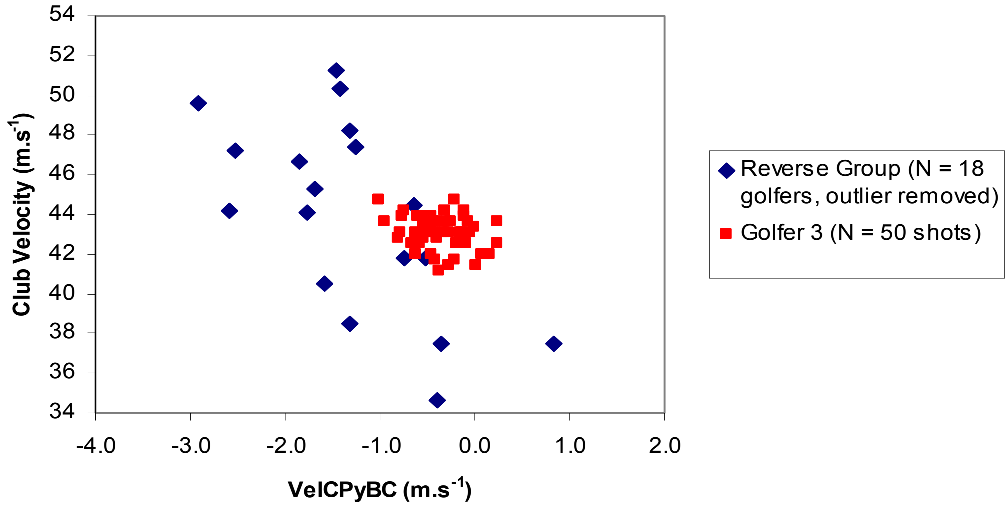


Figure 6.7: Scatterplot of VelCPyBC and Club Velocity for the Reverse group (*N* = 18, outlier removed) and Individual Golfer 3 (*N* = 50 shots).

6.4.5 Linear Relationships – general discussion

The finding that many relationships were linear on an individual basis requires further discussion. The relationship between Club Velocity and CPy% at swing events, maximum CPy, Minimum CPy and CPyR might be expected to be optimal rather than linear due to the limits imposed by stance width. Using CPyR as an example, four golfers indicated a linear relationship between CPyR and Club Velocity. This was appropriate conclusion based on the range of data produced by each golfer. However, to move CPy beyond the base of support will most likely lead to a performance decrement, possibly similar to the 'catastrophic' point reported by Best (1995) for javelin throwers (see figure 6.8). While requesting the golfer to produce a greater range of CPyR values could test this, this can affect ecological validity of the testing. It would be appropriate to indicate to golfers and coaches that the relationship between CPyR and Club Velocity was linear and that larger range of weight transfer was associated with better performance. However, it would also be necessary to detail the theoretical limit that exists but was not detected in this study and to detail the (obvious) issue of losing balance.

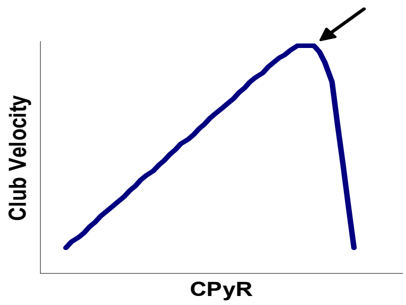


Figure 6.8: Example relationship between Club Velocity and CPyR showing a linear relationship up to a 'catastrophic failure point' denoted by an arrow

6.4.6 Statistical issues in single-subject designs

6.4.6.1 Generalisability

All individual golfers returned significant relationships between Club Velocity and CP parameters. More specifically, all golfers returned significant relationships between CPy% at swing events and Club Velocity and between CPy velocity at swing events and Club Velocity. Importantly, the single parameter CPyR was related to Club Velocity for all golfers. Given generalisability is based on replication, not sample size (Bates *et al.*, 2004) this would indicate strong support that the findings are generalisable on an individual basis. However, as these analyses still represent only five golfers of a population of over a million golfers in Australia (Australian Bureau of Statistics, 2001) the sample size was small. So while this study comprised five single-subject studies, and significant associations existed between CPy parameters and Club Velocity for all individual golfers, there is a need to further increase this number to assess if these relationships exist in the wider golfing population.

6.4.6.2 Normality

As this study examined curvilinear as well as linear relationships, the issue of normality was less influential to this analysis. Obviously for curvilinear relationships to exist, some deviation from normality must exist in at least one of the parameters being examined. As such, deviations from normality were a part of rather than a limitation of the analysis. Further, correlation (linear) analysis is robust to violations of normality (e.g. Pedhazur, 1997). These factors all support the issue of normality not affecting this study. However as there were a large number of linear relationships found and as non-normality has been a criticism of single subject designs (e.g. Reboussin and Morgan, 1996), the issue was explored.

The effects of any non-normal data on linear relationships were evaluated using within dataset distribution as outlined by Aron and Aron (1999). Repeating briefly, for a particular CP parameter-Club Velocity relationship, $N = 50$ randomly paired Club Velocity-CPy Parameter data pairs are assembled. The linear curve estimation is evaluated on this $N = 50$ dataset. This is repeated $N = 1000$ times to obtain 1000 R^2 values. These are sorted and the 975th ($p = 0.05$, two-tailed) R^2 value is determined. If the original R^2 value was greater than this value then this supports the relationship being a true rather than a random effect.

Within dataset distribution supported all linear results (table 6.20). For all relationships significant at $p < 0.05$, R^2 values were larger than the value indicated by the within dataset distributions as the threshold for significance at $p < 0.05$. This supported the results in this study and indicated that normality, or lack of, did not influence the analysis.

Table 6.20: Within dataset distribution for significant linear relationships between CP parameters and Club Velocity for individual golfers (N = 50 shots each).

		Original Correlation		Within dataset distribution	
		R^2	p	$p = 0.05$	$p = 0.01$
Golfer 1	CPy%LB	0.08	0.042	0.07	0.13
	VelCPyED	0.20	0.001	0.06	0.10
	VelCPyBC	0.13	0.012	0.07	0.12
	tVMaxCPy	0.08	0.045	0.07	0.12
	CPyR	0.20	0.001	0.05	0.18
Golfer 2	CPy%LB	0.30	<0.001	0.05	0.27
	CPy%ED	0.42	<0.001	0.05	0.32
	CPy%MD	0.21	0.001	0.05	0.13
	VelCPyLB	0.15	0.005	0.05	0.12
	VelCPyTB	0.28	<0.001	0.05	0.20
	VelCPyMD	0.34	<0.001	0.05	0.27
	VelCPyBC	0.11	0.020	0.05	0.12
	VMaxCPy	0.22	0.001	0.05	0.18
	MaxCPy%	0.36	<0.001	0.05	0.31
	MinCPy%	0.44	<0.001	0.04	0.30
	CPyR	0.52	<0.001	0.05	0.38
Golfer 3	CPy%MB	0.08	0.044	0.05	0.13
	CPy%LB	0.11	0.021	0.04	0.12
	CPy%TB	0.17	0.003	0.05	0.16
	CPy%ED	0.38	<0.001	0.05	0.35
	CPy%MD	0.50	<0.001	0.05	0.38
	CPy%BC	0.41	<0.001	0.05	0.34
	VelCPyTB	0.16	0.004	0.04	0.14
	VelCPyED	0.30	<0.001	0.05	0.23
	VMaxCPy	0.22	<0.001	0.04	0.19
	MinCPy%	0.11	0.020	0.06	0.12
Golfer 4	CPy%MB	0.12	0.014	0.05	0.12
	CPy%ED	0.10	0.027	0.04	0.10
	VelCPyMD	0.14	0.008	0.05	0.10
	CPyR	0.10	0.027	0.05	0.10
Golfer 5	VelCPyTB	0.27	<0.001	0.04	0.15
	tMinCPy%	0.15	0.005	0.04	0.12
	CPyR	0.08	0.042	0.05	0.10

Note: tMaxCPy for Golfers 2 and 3 have been eliminated from this assessment due to the identification of clusters detailed in section 6.4.2.1

6.4.6.3 Independence

Independence of samples was examined by calculating autocorrelations of residuals for each chosen relationship between the CP parameter and Club Velocity for three lags (three lags recommended by Bates *et al*., 2004). Lag 1 was the correlation between CP parameter N and CP parameter N+1 across the 50 shots. Lag 2 was the correlation between CP parameter N and CP parameter N+2 across the 50 shots. Lag 3 was the correlation between CP parameter N and CP parameter N+3 across the 50 shots. Tabachnick and Fidell (1996) and Hopkins (2003) recommended that these calculations should be performed on residuals. The threshold to indicate that a level of dependence existed that would influence regression statistics was set at R^2 = 0.50 as recommended by Bates *et al*. (2004). Table 6.21 reports this data for the Reverse and Front Foot golfers with CP parameters that were significantly associated with Club Velocity are shaded.

Table 6.21: Autocorrelations for residuals for 3 lags for Reverse golfers. Shaded areas indicate a significant relationship between the relevant CP parameter and Club Velocity

	Golfer 1 (Professional)			Golfer 2 (HCP = 3)			Golfer 3 (HCP = 14)			Golfer 4 (HCP = 5)			Golfer 5 (Social)		
Lag	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
CPy%TA	0.01	0.00	0.00	**0.10**	**0.10**	0.08	**0.19**	0.04	0.02	0.02	0.03	0.05	0.07	0.04	0.06
CPy%MB	0.02	0.02	0.03	0.07	0.01	0.04	**0.20**	0.04	0.03	0.01	0.07	0.02	0.07	0.05	0.07
CPy%LB	0.01	0.01	0.02	**0.22**	0.05	0.02	0.07	0.01	0.02	0.01	0.07	0.01	**0.08**	0.05	**0.08**
CPy%TB	0.03	0.04	0.04	**0.14**	0.04	**0.12**	0.07	0.00	0.00	0.02	0.04	0.04	**0.08**	0.06	0.06
CPy%ED	0.01	0.02	0.02	**0.22**	**0.14**	**0.12**	0.04	0.02	0.01	0.01	0.07	0.01	**0.08**	0.05	0.07
CPy%MD	0.00	0.00	0.00	0.04	0.01	0.02	0.02	0.02	0.00	0.03	0.03	0.03	0.04	0.01	0.04
CPy%BC	0.00	0.00	0.00	0.01	0.00	0.00	0.06	0.01	0.00	0.02	0.05	0.04	0.03	0.01	0.04
CPy%MF	0.01	0.01	0.01	0.01	0.01	0.04	**0.10**	0.02	0.00	0.02	0.03	0.05	0.03	0.00	0.03
VelCPyTA	0.00	0.00	0.00	0.04	0.02	0.00	**0.13**	0.02	0.03	0.01	0.04	0.02	0.04	0.02	0.02
VelCPyMB	0.00	0.00	0.01	0.02	0.00	0.04	**0.18**	0.03	0.02	0.02	0.04	0.05	0.06	0.04	0.06
VelCPyLB	0.01	0.01	0.01	0.02	0.02	0.00	**0.16**	0.01	0.00	0.01	0.05	0.03	0.06	0.04	0.06
VelCPyTB	0.02	0.02	0.02	**0.08**	0.05	0.06	**0.12**	0.01	0.01	0.02	0.02	0.04	0.02	0.02	0.00
VelCPyED	0.00	0.00	0.00	0.05	0.05	0.06	0.07	0.00	0.02	0.02	0.04	0.04	**0.08**	0.03	0.07
VelCPyMD	0.01	0.02	0.02	**0.19**	**0.12**	**0.12**	**0.12**	0.01	0.04	0.03	0.07	0.02	0.06	0.03	0.06
VelCPyBC	0.02	0.03	0.03	0.00	0.03	0.01	**0.10**	0.04	0.03	0.03	0.06	0.03	0.03	0.00	0.03
VelCPyMF	0.00	0.01	0.01	0.01	0.00	0.05	0.07	0.03	0.01	0.04	0.04	0.04	0.04	0.03	0.05
VMaxCPy	0.00	0.00	0.00	0.02	0.01	0.06	0.07	0.00	0.00	0.01	0.05	0.04	**0.08**	0.04	0.06
VtMaxCPy	0.03	0.03	0.03	0.03	0.11	0.02	**0.18**	0.03	0.03	0.02	0.06	0.03	0.07	0.03	0.06
MaxCPy%	0.00	0.00	0.00	**0.10**	0.03	0.05	0.04	0.00	0.00	0.02	0.07	0.03	0.05	0.01	0.04
tMaxCPy%	0.03	0.03	0.03	0.06	0.00	**0.19**	0.07	0.04	0.01	0.02	0.04	0.04	0.07	0.04	0.06
MinCPy%	0.02	0.03	0.03	0.04	0.02	0.01	0.02	0.00	0.00	0.01	0.04	0.03	**0.08**	0.07	0.04
tMinCPy%	0.01	0.01	0.01	0.22	0.23	0.07	**0.12**	0.01	0.00	0.02	0.05	0.04	0.03	0.05	0.03
CPyR	0.00	0.00	0.01	**0.13**	0.02	0.01	0.04	0.00	0.00	0.03	0.03	0.02	0.07	0.05	0.05

As no autocorrelation exceeded the threshold of $R^2 = 0.50$ (nor did any come close to this value), it was considered that the assumption of independence was not violated and use of regression statistics was appropriate for this data.

6.4.6.3.1 ISSUES ASSOCIATED WITH INDEPENDENCE

Three issues existed in the assessment of dependence of data in this study.

First is the threshold set to indicate dependence. While well below the threshold of $R^2 = 0.50$ suggested by Bates *et al.* (2004), a number of autocorrelations were significant at $p < 0.05$ in table 6.21 indicating a relationship, and hence a level of dependence existed in the data. This highlights that the decision on whether dependence exists in a dataset will be reliant on the threshold set by the researcher. There is little information in the literature to guide the researcher's decision as to what threshold constitutes dependence in a dataset. Hopkins (2003) suggested it is up to the researcher in each case to determine if dependence exists and if it is influential to the analysis. For this particular analysis, there were no relationships stronger than $R^2 > 0.25$ (half the threshold suggested by Bates *et al*. 2004) and so the decision to consider statistical results as valid is justified.

Second is what should be done about the analysis if dependence exists. The effect of dependent data will be to inflate R^2-values and reduce the range of confidence limits (Hopkins, 2003). As such, a possible approach where the researcher is concerned that dependence exists might be to make the alpha level more conservative. For this study, examining the relationships that were significantly autocorrelated at $p < 0.05$, the original relationship between Club Velocity and the relevant CP parametersreturned $p \leq 0.002$ for all but one relationship. As such, applying this more stringent statistical limitation would have changed the results of this study minimally.

The third is the use of autocorrelations to indicate dependence. This test was chosen as it was the only one used in the previous literature. However, it evaluates only the overall effects of the dataset. Dependence might exist in the short term (e.g. across 10 shots but not across all fifty shots) but not be detected by autocorrelations, which will evaluate all N, N+1 relationships. The use of different statistical tests to determine dependence is needed in future individual-based studies, although the issue as to what constitutes dependence and what should be done about it still remain. It should also be noted that some level of dependence can be an important finding in individual-based studies. This is the focus of the next section.

6.4.7 Non-linear analysis – Poincare plots

While single swings have been examined in golf, there has been little research into how performance alters between shots or over longer periods of time. The question of how performance changes from shot to shot and across shots over time is very important to the golfer. For example, is there a pattern of progression in performance leading up to a poor shot? Is there a pattern of performance after a poor shot is performed? Non-linear techniques, such as Poincare plots, could answer these questions. A Poincare plot is a nonlinear dynamic technique that plots a parameter value against its next value and this technique has been used to highlight nonlinear patterns in areas such as heart rate variability (e.g. Brennan *et al*., 2001; Kamen *et al*., 1996; Woo *et al*., 1992). As this study examined a large number of shots, it is worthwhile briefly examining this technique to indicate if it will provide useful information in the examination of weight transfer in the golf swing.

For each golfer, a Poincare plot of CPyR was generated with N on the x-axis and N+1 on the y-axis. CPyR was chosen as it was significantly associated to Club Velocity for all individual golfers. As well, referring to figure 6.9 for an explanation (and to Appendix F for proofs), the Poincare plots were quantified using

- R^2 values
- 95% ellipse area encompassed by the dot cloud using axes P1 and P2 (P1 is the line of identity where x = y. P2 is perpendicular to P1, figure 6.9)
- Short term variability (standard deviation of perpendicular distances from each [N,N+1] datapoint to P1)
- Long term variability (standard deviation of perpendicular distances from each [N,N+1] datapoint to P2)
- Histograms were calculated for Golfer 2 using perpendicular distances from P1 (line of identity) and P2 (axis perpendicular to line of identity).

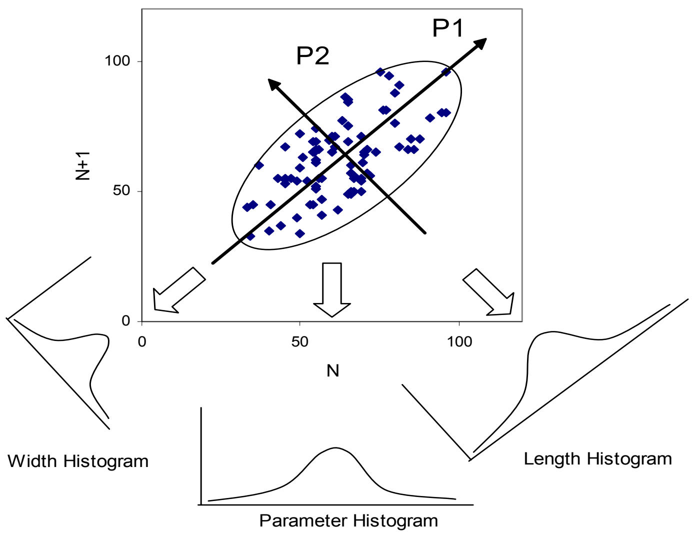


Figure 6.9: Quantification methods applied to the Poincare plot. P1 represents the line of identity (x = y). P2 is perpendicular to P1 passing through the mean of all values in the Poincare plot (adapted from Brennan *et al.*, 2001).

Figure 6.10 shows CPyR Poincare plots for each golfer.

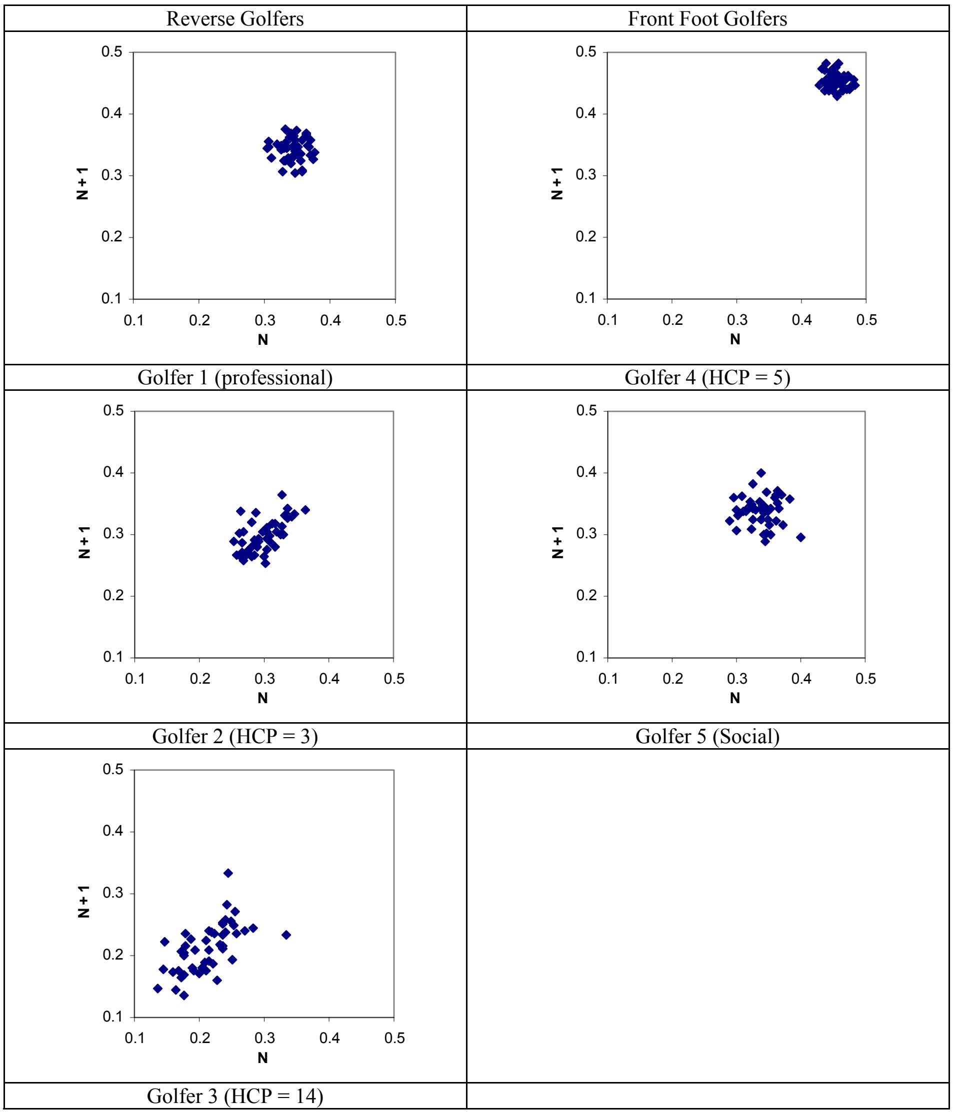


Figure 6.10: Poincare plot for CPyR values across 50 consecutive swings for all individual golfers.

Referring to figure 6.10, Golfer 1, Golfer 4 and Golfer 5 produced a rounded cloud while Golfer 2 and Golfer 3 produced a cloud that was elongated along the line of identity. The rounded clouds indicated similar variability existed in the short term (shot to shot) and long term (across *N* = 50 shots). However, the size of the point 'clouds' differed, indicating more variability for Golfer 5, less for Golfer 1 and less again for Golfer 4. Short and long term variability values and ellipse areas presented in table 6.22 support both observations. The elongated clouds of Golfer 2 and Golfer 3 indicated more long term compared with short term variability (also supported by data in table 6.22).

Table 6.22: Ellipse area, short term variability and long term variability for CPyR across 50 swings for all individual golfers

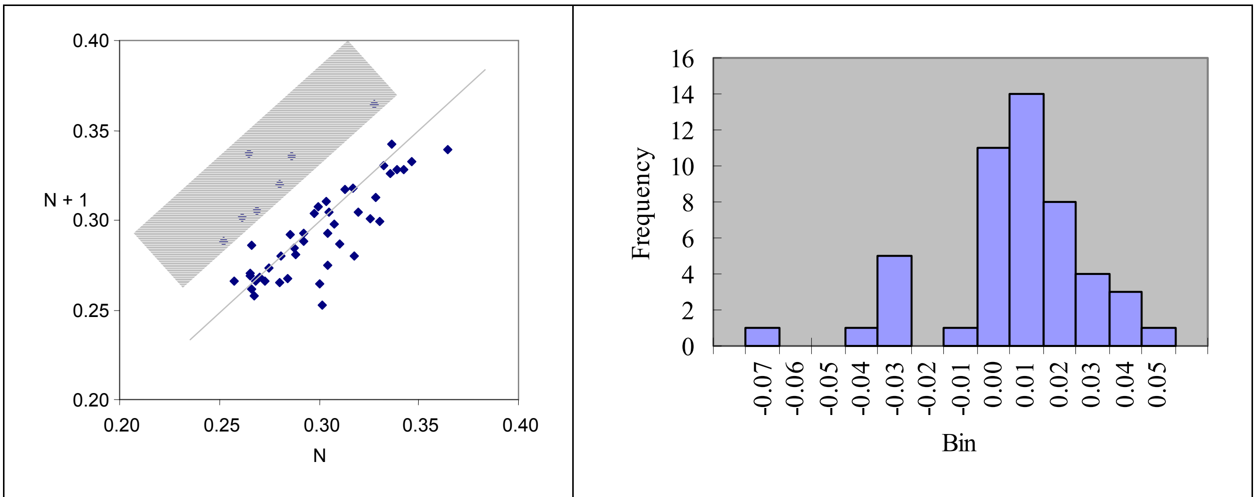
	Golfer 1 (Professional)	Golfer 2 (HCP = 2)	Golfer 3 (HCP = 14)	Golfer 3 (HCP = 14)	Golfer 4 (HCP = 5)	Golfer 5 (Social)
				minus outlier		
R^2	0.00	0.40	0.37	0.42	0.00	0.00
p	0.836	<0.001	<0.001	<0.001	0.474	0.831
Ellipse 95% (m^2)	0.015	0.029	0.061	0.037	0.008	0.029
Short Term variability (m)	0.017	0.016	0.024	0.021	0.012	0.024
Long Term variability (m)	0.018	0.035	0.050	0.035	0.013	0.025

$N = 49$ for all analyses except Golfer 3 outlier removed ($N = 47$)
Note: For outlier removal in Poincare plots, the two points that the outlier was linked to were removed i.e.[$N_{outlier}$, $N_{outlier}$ + 1] and [N, $N_{outlier}$], (where $N_{outlier}$ is N+1).

Two golfers produced a significant R^2 value (Golfer 2 and Golfer 3). For both, the relationships were positive (refer figure 6.10) indicating large CPyR values tended to follow large CPyR values and small CPyR values followed small CPyR values. For this to exist there must be long term drift. That is, larger values tended to follow larger values but over time these values drifted such that smaller values were produced together. This is distinct to short term changes that will produce a negative correlation (small values followed by large values represented in Poincare plots by a bottom right to top left spread of points). No relationship existed for the remaining three golfers.

Poincare plot analysis provided interesting short term shot to shot patterning information for Golfer 2. The jumps from relatively smaller to larger CPyR values tended to be large while from large to small CPyR values tended to be more evenly distributed. This was evident in the Poincare plot and in the width histogram, which showed a group of seven points that were distinct from the remaining datapoints

(Poincare plot presented again in figure 6.11 i, with the seven points denoted by a shaded area). Examining the distribution of the difference in CPyR between shots ($CPyR_N$-$CPyR_{N+1}$, figure 6.11 ii) it can be seen that the same seven points differed from the next swing by more than 0.03 m with all but one of the remaining swing to swing differences where CPyR moved from small to large were less than 0.01 m.



i. Poincare Plot	ii. Difference in CPyR between shots Histogram

Figure 6.11: Poincare plot and histogram of the difference between consecutive shots for CPyR [N-(N+1)] across 50 swings for Golfer 2. Negative values mean the next swing produced a larger CPyR value. The bin numbers represent the upper value of the range.

The underlying cause for the difference in small to large CPyR changes compared to large to small CPyR changes from shot to shot is not clear. In all cases for the seven swings, Club Velocity also increased for the corresponding swing (on average by 1.5 $m.s^{-1}$ compared with the mean change between all other shots of 0.8 $m.s^{-1}$ [absolute change values used for this calculation]). It could be a conscious effort by the golfer to increase this range (or Club Velocity) after a smaller than desired CPyR (or Club Velocity) for the previous swing. However, no data was collected to indicate if this golfer was making conscious control or change. As such, there is a need to gather information on each trial for use in a shot to shot analysis. This should be considered in designs of any future study using Poincare plot in analysis of the golf swing.

While there has been criticism that the measures used to assess Poincare plots provide no more information than linear measures (e.g. Brennan *et al*., 2001), short and long term variability measures provided information not evident from linear measures in this analysis. For example, the linear measure standard deviation (i.e. original or overall SD) for Golfer 2 was considerably larger that for Golfer 1 (table 6.23). However, short term variability for Golfer 2 was actually smaller than that for Golfer 1, with the larger linear standard deviation a function of the increased long term variability. This indicated that from shot to shot, Golfer 2 was slightly more consistent than Golfer 1, but over a longer period, CPyR performance drifted. A similar contrast existed between Golfer 3 and Golfer 5. This information was not evident with linear measures and so the non-linear measures have provided more information.

Table 6.23: Comparison of linear (standard deviation) and non-linear (Poincare plot measures) indices for CPyR for individual golfers.

	Values				Rank			
	Linear Measure	Poincare Plot Measures			Linear Measure	Poincare Plot Measures		
	SD	ST	LT	E95%	SD	ST	LT	E95%
Golfer 1	0.018	0.018	0.018	0.016	2	3	2	2
Golfer 2	0.027	0.016	0.033	0.030	4	2	4	4
Golfer 3	0.035	0.021	0.035	0.037	5	5	5	5
Golfer 4	0.013	0.012	0.011	0.012	1	1	1	1
Golfer 5	0.024	0.023	0.023	0.027	3	4	3	3

SD = standard deviation (m)
ST = short term variability (m)
LT = long term variability (m)
E95% = area of 95% ellipse (m^2)
* Outliers removed for Golfer 3 - no change to rankings with or without Golfer 3 outlier

The ellipse measure did not provide more information for this analysis than that provided by the linear standard deviation. The ranking of golfers from smallest to largest SD was the same as for the ellipse area. As well, like linear standard deviation, the ellipse measure was not sensitive to differences in short and long term variability. For example, Golfer 2 and Golfer 5 produced the same ellipse area but distinctly different axis lengths (i.e. short and long term variability). It would seem that this measure is not required as short and long term variability measures provide more useful information, and this is further emphasized with linear standard deviation providing the same relative information as the ellipse measure.

Based on the interesting finding for Golfer 2 and the useful information provided by separating short and long term variability, it would be worth continuing to explore Poincare plots in future research. These need to include examining technical aspects 'on site' where the actual result of the swing (distance and accuracy) can be determined. As well, a larger number of shots would be useful in analysis, although as this would be required to be performed over a period of time, conscious technique changes and day to day variability issues would need to be addressed. Other techniques such as detrended fluctuation analysis which looks at short and long term patterns across a large *N* (e.g. Hausdorff *et al*., 2001) may also hold useful information.

6.5 CONCLUSION

Weight transfer is important to performance of the golf swing on an individual basis. All individual golfers returned significant relationships between CP parameters and Club Velocity. However the nature of this relationship is individual-specific with different parameters associated with Club Velocity for different golfers. As well, effect sizes, and in some cases the direction of relationships, differed between golfers for the same CP parameter. This is the first statistical support for the importance of weight transfer in the golf swing on an individual basis.

Both positioning of weight and rate of weight transfer at specific swing events was important to all golfers tested. All golfers returned significant relationships between Club Velocity and CPy% at swing events. As well, all golfers returned significant relationships between Club Velocity and CPy velocity at swing events. However, the swing events indicated as important differed between golfers, as did effect sizes and the nature of these relationships (linear, polynomial, cubic). This indicated a complex and individual-specific relationship between both positioning of weight and rate of weight transfer with Club Velocity.

Weight transfer range and position of weight at the late backswing event were the CPy parameters most often related to Club Velocity. All golfers returned a significant association between CPy range and Club Velocity. For four golfers the indicated relationship was linear with a larger weight transfer range associated with a larger Club Velocity at ball contact and on a general level, the other golfer also produced this relationship. Four golfers returned a significant relationship between CPy% at late backswing and Club Velocity. For three golfers a smaller CPy%, or positioning CPy% nearer the back foot at late backswing was associated with larger Club Velocities. On a general level, this relationship also existed for the fourth golfer. The importance of this swing event might be related to its proximity to the start of the forces associated with moving the club forward. These results supported the importance of weight transfer range in the golf swing and highlighted the need to include the late backswing event in future studies.

For Reverse golfers, both CPy%LB and VelCPyBC were important on an individual-basis. For all three Reverse golfers a significant negative linear relationship indicated a position further towards the back foot at LB was associated with a larger Club Velocity. As well, all three Reverse golfers produced at least a small negative effect for the relationship between VelCPyBC and Club Velocity indicating moving CPy towards the back foot more rapidly at BC was associated with a larger Club Velocity. That weight is moving towards the back foot and a more rapid movement towards the back foot is of advantage is a new finding that has not been identified or discussed in either coaching or scientific literature.

Non-linear techniques such as Poincare plots hold useful information in assessment of weight transfer in the golf swing. Differences existed in short term (shot to shot) and long term (across all fifty shots) variability both within and between golfers. As well, for one golfer, the shot to shot changes in range of weight transfer differed depending on whether the change was to increase weight range in the next shot, or to decrease it. This information was not evident using linear measures such as overall standard deviation.

A number of important future directions exist for examining weight transfer in the golf swing on an individual basis. The number of subjects examined needs to increase, as does the number of shots performed to allow for the examination of long term patterns using techniques such as detrended fluctuation analysis. Further, the inclusion of kinematic data is essential in better defining the underlying mechanics generating the significant findings in this study between weight transfer and Club Velocity.

CHAPTER 7: GENERAL CONCLUSIONS

Different weight transfer styles exist in the golf swing. Different styles exist for CPy% movement for golfers across skill levels. In this study, two major groups were identified; a Front Foot group and a Reverse group. For both groups, CPy% was positioned near the midpoint of stance at takeaway, near the back foot at mid, late and the top of backswing, before moving towards the front foot at early downswing. The Front Foot group continued to move CPy% towards the front foot, with CPy% being positioned close to the front foot at ball contact and mid follow through. However, for the Reverse group, CPy% moved towards the back foot again and was positioned at the midpoint of the feet at ball contact and closer to the back foot at mid follow through. As both groups contained highly skilled golfers and no difference existed between handicap or club velocity at ball contact between the two groups, neither style was deemed to be a technical error.

Different styles need to be identified prior to any group-based analysis. Different parameters were important for the Front Foot group and the Reverse group in this study. Further, errors would have been made if the groups had been treated as one. Weight transfer range was significant on a group basis for the whole group but this would have represented a type 2 error for the Reverse group, for which this parameter was not significant. As well, weight position and rate of weight transfer, significant for the reverse group, were not significant when the groups were analysed together, producing a type 1 error for the Reverse group for these parameters.

A new and important finding was that for Reverse golfers, at ball contact weight was positioned near midstance and was moving towards the back foot. Further, moving weight more rapidly towards the back foot at this event was related to better performance on a group and an individual basis. This is in direct conflict with the coaching literature which reports weight should be nearer the front foot at ball contact and be continuing to move further towards the front foot. There is no mention of this technique in either coaching or scientific literature so it represents a new finding. This finding requires that both coaching and scientific recommendations for weight transfer in the golf swing need to be changed.

Weight transfer is important to performance on a group basis for both the Front Foot and Reverse styles. For the Front Foot style, an increased range of CPy movement and an increased rate of CPy movement towards the front foot in downswing were associated with larger club velocities at ball contact. For the Reverse style, positioning CPy nearer midstance (compared to further towards the back foot) and a larger rate of CPy movement towards the back foot at ball contact were associated with larger club velocities.

Weight transfer is important on an individual basis for Front Foot and Reverse style golfers. Positioning weight during the swing as well as the rate of weight transfer were both important, with all golfers returning at least one significant result for both. Range of weight transfer was important for all golfers and on a general level, a larger range of weight transfer was associated with a larger club velocity and ball contact. As well, for four golfers, positioning of weight at the late backswing event was related to performance indicating this event is important. However, results were individual specific with each golfer returning a unique set of significant results.

Both individual-based and group-based analyses are required to extract all the useful information from data. In this thesis, important information was evident on a group basis that was not evident on an individual basis. Similarly, important information was found on an individual basis that would have been missed if only group-based assessment was used. The use of both types of analysis offered a more thorough and useful investigation of the importance of weight transfer in the golf swing.

The use of more swing events than has been used in previous research is essential in weight transfer research. Using only the most common events of takeaway, top of backswing and ball contact, important relationships would have been missed on both a group and an individual basis. Importantly, the swing event at late backswing was important for four of five golfers as well as important for Reverse golfers on a group basis. Future work must include this swing event.

The use of more trials per individual is necessary to obtain stable mean parameter values for research into weight transfer in the golf swing. As few as three and up to ten trials were required for mean CP parameter values to stabilise for different individual golfers. Only one previous study has used ten trials to establish mean values for individual subjects for use in statistical analysis.

Future work requires similar analysis with larger *N* for both group-based and individual-based analyses. On a group basis, increasing *N* will provide a more powerful analysis to determine if more styles exist. On an individual basis, a better profile of important parameters can be obtained. This work needs to be expanded to include more clubs such as fairway woods and irons to identify if styles are consistent across all clubs or if these styles are distinctive to the driver. As part of this work kinematic analyses should be included to explore the mechanics underlying the two swing styles, and any other style that might be evident. As well, including CPx data in the analysis (perpendicular to the line of shot) will provide a more thorough assessment of forces at the feet during the swing.

A number of statistical and analytical tools that have not been applied to the analysis of the golf swing should be explored. In this study, a non-linear technique identified important information for some golfers. Further development of this work is important examining both short and long term patterns of performance. The use of detrended fluctuation analysis can identify if long-term patterns occur across a large number of shots. Neural networks and fuzzy clustering should be explored to identify if these methods add more useful information to the cluster analysis and improve the process (e.g. speed classification up) of the formal classification of golfers.

REFERENCES

Aldenderfer, M.S. and Blashfield, R.K. (1984). *Cluster analysis*. California: Sage Publications.

AMTI (1982). Computerised biomechanics platform system manual. Advanced Medcal Technology Inc, Massachusetts.

Aron, A. and Aron, E.N. (1999). *Statistics for psychology (2nd Edition).* New Jersey: Prentice Hall.

Arutyunyan, G.H., Gurfinkwl, V.S. and Mirskii, M.L. (1968). Investigation of aiming at a target. *Biophysics*, **14**, 1162-1167.

Australian Bureau of Statistics (1999). *Sport attendance, Australia 1999*. Catalogue No. 4174.0. Australian Bureau of Statistics.

Baker, F.B. and Hubert, L.J. (1972). Measuring the power of hierarchical cluster analysis. *Journal of the American Statistical Association*, **70**, 31-38.

Ball, K. (1999). *Body sway and aim point fluctuation in rifle and pistol shooting*. Unpublished Masters Thesis. Victoria University, Australia.

Ball, K., Best, R. and Wrigley, T. (2001). Comparison of three methods for determining the optimal cut-off frequency of a Butterworth filter. In *Proceedings of the XVIIIth Congress of the International Society of Biomechanics* (Edited by Gerber H. and Muller, R.) p 40. ETH Zurich: Interrepro AG Münchenstein.

Ball, K., Best, R. and Wrigley, T. (2003a). Body sway, aim point fluctuation and performance in rifle shooting: inter and intra-individual analysis. *Journal of Sports Sciences,* **21**, 559-566.

Ball, K., Best, R. and Wrigley, T. (2003b). Inter and intra-individual analysis in elite sport: pistol shooting. *Journal of Applied Biomechanics,* **19**, 28-38.

Ball, K., Best, R., Dowlan, S. and Brown, D. (2002). Different weight transfer patterns in golf. In *Proceedings of the XXth International Symposium of Biomechanics in Sports* (Edited by Giannikellis, K.), pp 192-195. University of Extremadura, Spain: Servico do Publicaciones

Barlow, D. H. and Hersen, M. (1973). Single-case experimental design. *Archive of General Psychiatry*, **29**, 319-325.

Barrentine, S.W., Flesig, G.S. and Johnson, H. (1994). Ground reaction forces and torques of professional and amateur golfers. In *Science and Golf II: Proceedings of the 1994 World Scientific Congress of Golf* (Edited by Cochran, A.J. and Farrally, F.R.), pp 33-39. London: E & FN Spon.

Bates, B.T. (1996). Single-subject methodology: an alternative approach. *Medicine and Science in Sports and Exercise,* **28**, 631-638.

Bates, B.T. and Stergiou, N. (1996). Performance accommodation to mid-sole hardness during running. *Journal of Human Movement Studies*, **21**, 189-210.

Bates, B.T., James C.R. and Dufek, J.S. (2004). Single-subject analysis. In *Innovative Analyses of Human Movement* (Edited by Stergiou, N), pp 3-28. Champaign, Illinois: Human Kinetics.

Bates, B.T., James, S.L. Osternig, L.R. and Mason, B.R. (1979). Functional variability of the lower extremity during the support phase of running. *Medicine and Science in Sports*, **11**, 328-331.

Bates, B.T., Osternig, L.R., Sawhill, J.A. and James, S.L. (1983). An assessment of subject variability, subject-shoe interaction, and the evaluation of running shoes using ground reaction force data. *Journal of Biomechanics,* **16**, 181-191.

Begg, R.K., Wytch, R. and Major, R.E. (1990). A microcomputer-based video vector system for clinical gait analysis. *Journal of Biomedical Engineering,* **12**, 383-387.

Best, R. (1995). Performance optimisation in sport: results of computer simulations. In *Proceedings of the Australian Conference of Science and Medicine in Sport*, pp 56-57. Canberra: Sports Medicine Australia.

Best, R.J., Begg, R.K., Ball, K.A. and James, L. (2000). Minimum foot clearance variability in walking: a case study. In *Book of Abstracts of the 3rd Australasian Biomechanics Conference* (Edited by Barrett, R., Simeoni, R., & D'Helon, C.), pp 7-8. Gold Coast, Australia: Griffith University.

Bobbert, M. F. and Schamhardt, H. C. (1990). Accuracy of determining the point of force application with piezoelectric force plates. *Journal of Biomechanics,* **23,** 705-710.

Boesch, D.F. (1977). *Application of numerical classification in ecological investigations.* Corvallis, OR: U.S.E.P.A.

Breckenridge, J.N. (1989). Replicating cluster analysis: method, consistency and validity. *Multivariate Behavioral Research*, **24**, 147-161.

Brennan, M., Palaniswami, M. and Kamen, P. (2001). Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE transactions on Biomedical Engineering*, **48**, 1342-1347.

Brown, D., Best, R., Ball, K. and Dowlan, S. (2002). Age, centre of pressure and clubhead speed in golf. In *Proceedings of Science and Golf V* (Edited by Thain, E.), pp 40-53. London:Routledge.

Browne, D. (2002). *Weight transfer patterns between different skill levels and clubs in golf*. Unpublished Masters Thesis. Victoria University, Australia.

Burden, A.M., Grimshaw, P.N. and Wallace, E.S. (1998). Hip and shoulder rotations during the golf swing of sub-10 handicap players. *Journal of Sports Sciences*, **16**, 165-176.

Carlsoo, S. (1967). A kinetic analysis of the golf swing. J*ournal of Sports Medicine and Physical Fitness,* **7**, 76-82.

Caster, B.L. and Bates, B.T. (1995). The assessment of mechanical and neuromuscular response strategies during landing. *Medicine and Science in Sports and Exercise,* **27**, 736-744.

Challis, J.H. (1999). A procedure for the automatic determination of filter cutoff frequency for the processing of biomechanical data. *Journal of Applied Biomechanics,* **15***,* 303-317.

Chau, T. (2001). A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods. *Gait and Posture*, **13**, 49-66.

Cochran, A. and Stobbs, J. (1968). *Search for the perfect swing*. Golf Society of Great Britain. Heinemann: London.

Cohen, J. (1988). *Statistical power analysis for the behavioural Sciences (2nd Edition).* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Coleman, S. (1999). *Biomechanics and its application to coaching*. In The Coaching Process (Edited by Cross, N., and Lyle, J.), pp 53-78. Oxford: Butterworth-Heinemann.

Cooke B (1987). *Golf: The professional approach*. Twickenham, England: Hamilton Publishing Group.

Cooper, J.M., Bates, B.T. Bedi, J. and Scheuchenzuber, J. (1974). Kinematic and kinetic analysis of the golf swing. In *Biomechanics IV* (Edited by Nelson, R.C. and Morehouse, C.A.), pp 298-305. Baltimore: University Park Press.

Daniel, C. and Wood, F.S. (1980). *Fitting equations to data (2nd Edition).* New York: John Wiley & Sons.

Davids, K., Glazier, P., Araújo, D. and Bartlett, R. (2003). Movement systems as dynamical systems: the functional role of variability and its implications for sports medicine. *Sports Medicine*, **33**, 245-260.

Dowlan, S., Brown, D., Ball, K., Best, R.J. and Wrigley, T. (2001). Intra-subject variability in the golf swing: a comparison of high and low handicap players In *Proceedings of the XIXth International Symposium of Biomechanics in Sports* (Edited by Blackwell, J.) p 120. University of San Francisco, USA: University of San Francisco.

Dufek, J.S., Bates, B.T., Stergiou, N. and James, C.R. (1995). Interactive effects between group and single subject response pattern. *Human Movement Science,* **14**, 301-323.

Elliot, B., Marsh, T. and Blanksby, B. (1986). A three dimensional cinematographic analysis of the tennis serve. *International Journal of Sports Biomechanics*, **2**, 260-271.

Everitt, B.S. (1979). Unresolved problems in cluster analysis. *Biometrics*, **35**, 169-181.

Forner-Cordero, A., Koopman, H.J.F.M. and van der Helm, F.C.T. (2006). Describing gait as a sequence of states, *Journal of Biomechanics*, **39**, 948-957.

Forwood, M.R., Clarke, D. and Wilson, B.D. (1985). A cinematographic analysis of the front handspring to front somersault. *The Australian Journal of Science and Medicine in Sport*, **17**, 8-14.

Frank, J. (1994). *The weight shift board*. Golf Magazine, **3**, 152.

Giakas, G. and Baltzopoulos, V. (1997). Optimal digital filtering requires a different cut-off frequency strategy for the determination of the higher derivatives. *Journal of Biomechanics,* **30**, 851-855.

Gower, J.C. (1975). Goodness-of-fit criteria for classification and other patterned structures. In *Proceedings of the 8th International Conference on Numerical Taxonomy* (Edited by Estabroo, G), pp. 38-62. San Francisco: Freeman.

Grabe, S.A. and Widule, C.J. (1988). Comparative biomechanics of the jerk in Olympic weightlifting. *Research Quarterly for Exercise and Sport*, **59**, 1-8.

Grant, T., Bann, S. and Lynch, D. (1996). *Play like the pros*. Melbourne: Wilkinson Books.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995). *Multivariate data analysis with readings (4th Edition).* Englewood Cliffs, New Jersey: Prentice-Hall.

Hausdorff, J. M., Nelson, M. E., Kaliton, D., Layne, J. E., Bernstein, M. J., Nuernberger, A. and Faitarone Singh, M. A. (2001). Etiology and modification of gait instability in older adults: a randomized controlled trial of exercise. *Journal of Applied Physiology,* **90**, 2117-2129.

Hayes, G.J. (1970). *Numerical approximation to functions and data*. University of London: The Athlone Press.

Hodge, K. and Petlichkoff, L. (2000). Goal profiles in sport motivation: a cluster analysis. *Journal of Sport and Exercise Psychology*, **22**, 256-272.

Holgersson, M. (1978). Multivariate stratification with the use of cluster analysis. *Journal of the American Statistical Association*, **73**, 397-403.

Hopkins, W. (1999). *Planning, performing and publishing research with confidence limits*. Sportsci.org (accessible at: http://sportsci.org/resource/stats/errors.html#inflation). Viewed 10 March 2001).

Hopkins, W. (2001). *Statistical models: Independence of Observations* (accessible at: http://sportsci.org/resource/stats/modelsdetail.html#residuals). Viewed 2 March 2001.

Hopkins, W. (2002). *A Scale of Magnitudes for Effect Statistics* (accessible at: http://sportsci.org/resource/stats/effectmag.html). Viewed 25 July 2002.

Hubert L.J. and Levin J.R. (1976). A general statistical framework for assessing categorical clustering in free recall. *Psychological Bulletin*, **83**, 1072-1080.

Kamen, P.W., Krum, H. and Tonkin, A.M. (1996). Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity. *Clinical Science*, **91**, 201-208.

Kelley, B. (2006). *Golf club distances*. Beginners Guide to Golf (accessible at: htm://golf.about.com/od/beginners/i/blclubdistance.htm). Viewed 20 March 2006.

Knudson, D. and Bahamonde, R. (2001). Effect of endpoint conditions on position and velocity near impact in tennis. *Journal of Sports Sciences,* **19**, 839-884.

Koenig, G., Tamres, M. and Mann, R.W. (1993). An analysis of the kinetics and kinematics of the golf swing. In *Biomechanics in Sport XI: proceedings of the XIth Symposium of the International Society of Biomechanics in Sports* (Edited by Hamill, J.), pp. 328-333. Amherst, Massachusetts: University of Massachusetts.

Kos, A.J. and Psenick, C. (2000). Measuring cluster similarity across methods. *Psychological Reports*, **86**, 858-862.

Koslow, R. (1994). Patterns of weight shift in the swings of beginning golfers. *Perceptual and Motor Skills*, **79**, 1296-1298.

Kreighbaum, E. and Barthels, K.M. (1985). Biom*echanics: a qualitative approach for studying human movement*. Minneapolis: Burgess Publishing.

Lafond, D., Duarte, M. and Prince, F. (2004). Comparison of three methods to estimate the centre of mass during balance assessment. *Journal of Biomechanics*, **37**, 1421-1426.

Leadbetter, D. (1993). *Faults and fixes: The golfer's essential problem-solving manual*. London: Collins Willow.

Leadbetter, D. (1995). *A Lesson with Leadbetter: The swing*. London, United Kingdom: Telstar video entertainment.

Leaderboard (2006). *Distance table*. Leaderboard.com (accessible at: http://www.leaderboard.com/glossary_distancetable). Viewed 20 March 2006.

Lockwood, J. 1998. A small-scale local survey of age-related male golfing ability. In *Science and Golf III. Proceedings of the World Scientific Congress of Golf* (Edited by Cochran, A. and Farrally, M.), pp. 112-119. United Kingdom: Human Kinetics.

Madonna B (2001). *Coaching golf successfully*. Champaign, Illinois: Human Kinetics.

Mason, B.R., McGann, B. and Herbert, R. (1995). Biomechanical golf swing analysis. In *Proceedings of the XIII International Symposium for Biomechanics in Sport* (Edited by Bauer, T.), pp. 67-70. Thunder Bay, Ontario: Lakehead University.

Mason, B.R., Thinnes, R. and Limon, S. (1990). A kinetic and kinematic analysis of the golf swing. In *XIII International Congress in Biomechanics: Book of Abstracts* (Edited by Marshall, R.N., Wood, G.A., Elliott, B.C., Ackland, T.R. and McNair, P.J.), pp. 47-48. Perth, Australia: University of Western Australia.

McIntyre, R.M. and Blashfield, R.K. (1980). A nearest-centroid technique for evaluating the minimum-variance clustering procedure. *Mutlivariate Behavioural Research*, **15**, 225-238.

Meadows, C. (2001). *The encyclopedia of golf techniques*. Bath, United Kingdom: Paragon.

Meinen, L. (2006). *Golf club distance chart*. Bellaonline (accessible at: http://www.bellaonline.com/articles/art3336.asp). Viewed 20 March 2006.

Middleton, J., Sinclair, P. and Patton, R. (1999). Accuracy of centre of pressure measurement using a piezo-electric force platform. *Clinical Biomechanics*, **14**, 357-360.

Milligan, G.W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, **45**, 325-343.

Milligan, G.W. (1981). A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Multivariate Behavioural Research,* **16**, 379-407.

Milligan, G.W. (1996). Clustering validation: results and implications for applied analysis. In *Clustering and Classification* (Edited by Arabie, P., Hubert, L.J. and De Soete, G.), pp 341-376. River Edge, New Jersey: World Scientific Publishing.

Milligan, G.W. and Cooper, M. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, **50**, 159-179.

Morey, L. and Agresti, A. (1984). The measurement of classification agreement: an adjustment to the Rand statistic for chance agreement. *Education and Psychological Measurement*, **44**, 33-37.

Mullineaux, D. R., Bartlett, R. M. and Bennet, S. (2000). Research design and statistics in biomechanics and motor control. *Journal of Sports Sciences*, **19**, 739-760.

Nagano, N. and Sawada, Y. (1977). A kinematic analysis of the golf swing by means of fast motion pictures in connection with wrist action. *Journal of Sports Medicine,* **17**, 413-419.

Nakamura, S. (1993). *Applied numerical methods in C*. New Jersey: Prentice Hall.

Neal, R. (1998). Golf swing styles: a kinetic and 3D kinematic comparison. *Australian Conference of Science and Medicine in Sport* (accessible at: www.ausport.gov.au/fulltext/1998/acsm/smabs183.html). Viewed 10 November 2001.

Newell S and Foston P (1995). *How to play golf.* London: Annes Publishing.

Norman G (1995). *Advanced golf.* Port Melbourne, Australia: William Heinemann.
Pedhazur, E. J. (1997). *Multiple regression in behavioural research*. Fort Worth, Texas: Harcourt Brace College Publishers.

Putnam, C. (1993). Sequential motions of the body segments in striking and throwing skills: descriptions and explanations. *Journal of Biomechanics*, **26** (suppl. 1), 125-136.

Quintavala, S.J. (2006). Experimental determination of the effects of clubhead speed on driver launch conditions and the effects on drive distance for balls used by the PGA tour. USGA Technical Report RB/cor2006-01.

Rae, I., Fairweather, M. and Sanders, R. (2001). *Using your head to improve your golf swing.* Coaches Information Service, (accessible at: http://www.coachesinfo.com/category/golf/58/). Viewed 10 November, 2001.

Reboussin, D.M. and Morgan, T.M. (1996). Statistical considerations in the use and analysis of single-subject design. *Medicine and Science in Sports and Exercise,* **28**, 639-644.

Richards, J., Farrell, M., Kent, J. and Kraft, R. (1985). Weight transfer patterns during the golf swing. *Research Quarterly for Exercise and Sport,* **56**, 361-365.

Robinson, R.L. (1994). A study of the correlation between swing characteristics and clubhead velocity. In *Science and Golf II: Proceedings of the 1994 World Scientific Congress of Golf* (Edited by Cochran, A.J. and Farrally, F.R.), pp 84-90. London: E & FN Spon.

Rohlf, F.J. (1974). Methods of comparing classifications. *Annual Review of Ecology and Systematics*, **5**, 101-113.

Salo, A.I.T., Grimshaw, P.N. and Viitasalo, J.T. (1997). Reliability of variables in the kinematic analysis of sprint hurdles. *Medicine and Science in Sports and Exercise,* **29**, 383-389.

Schmeidmayer, H. and Kastner, J. (2000). Enhancements in the accuracy of the centre of pressure (COP) determined with piezo-electric force plates are dependent on load distribution. *Journal of Biomechanical Engineering,* **122**, 523-527.

Schollhorn, W.I., Nigg, B.M., Stefanyshyn, D.J. and Liu, W. (2002). Identification of individual walking patterns using time discreet and time continuous data sets. *Gait and Posture*, **15**, 180-186.

Sommer, R., Kohler, D. and Calame, C. (1997). Centre of pressure (COP) enhancement in piezoelectric force plates. *Proceedings of the XVIth Congress of the International Society of Biomechanics (Edited by),* p 18.Tokyo: University of Tokyo.

SPSS 10.0 Guide (1997). SPSS Inc: Chicago.

Suttie, J. (2006). *Your perfect swing*. Champaign, Illinois: Human Kinetics.

Tabachnick, B.G. and Fidell, L.S. (1996). *Using multivariate statistics (3rd Edition).* New York: Harper-Collins.

Taylor, J.R. (1982). *Introduction to error analysis: the study of uncertainties in physical measurements (2nd edition).* Mill Valley: University Science Books.

Thorndike, R.L. (1953). Who belongs in a family? *Psychometrika*, **18**, 267-276.

Vardaxis, V.G., Allard, P., Lachance, R. and Duhaimme, M. (1998). Classification of able bodied gait using 3-D muscle powers. *Human Movement Science,* **17**, 121-136.

Wallace, E.S., Graham, D. and Bleakley, E.W. (1990). Foot to ground pressure patterns during the golf drive: a case study involving a low handicap player and a high handicap player. In *Science and Golf: Proceedings of the First World Scientific Congress of Golf* (Edited by Cochran, A.J.), pp. 25-29. London: E & FN Spon.

Wallace, E.S., Grimshaw, P.N. and Ashford, R.L. (1994). Discrete pressure profiles of the feet and weight transfer patterns during the golf swing. In *Science and golf II: proceedings of the 1994 World Scientific Congress of Golf* (Edited by Cochran, A.J. and Farrally, F.R.) pp 26-32. London, E & FN Spon.

Wang, E., Miller, L.D., Ohnmacht, G.A., Mocellin, S., Perez-Diez, A., Petersen, D., Zhao, Y., Simon, R., Powell, J.I., Asaki, E., Alexander, H.R., Duray, P.H., Herlyn, M., Restifo, N.P., Liu, E.T., Rosenberg, S.A. and Marincola, F.M. (2002). Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. *Cancer Research,* **62**, 3581-3586.

Williams, K.R. and Cavanagh, P.R. (1983). The mechanics of foot action during the golf swing and implications for shoe design. *Medicine and Science in Sport and Exercise,* **15**, 247-255.

Wilson, B.D. and Howard, A. (1983). Movement pattern recognition in description and classification of backstroke swim start. *Journal of Human Movement Studies,* **9**, 71-80.

Winter, D. A. (1990) *Biomechanics and motor control of human movement (2nd Edition).* New Jersey: Wiley.

Winter, D.A. (1995). *ABC (anatomy, biomechanics and control) of balance during standing and walking*. Waterloo: University of Waterloo.

Woo, M.A., Stevenson, W.G., Moser, D.K., Trelease, R.B. and Harper, R.H. (1992). Patterns of beat to beat heart rate variability in advanced heart failure. *American Heart Journal*, **123**, 704-710.

Yessis M (2000). *Explosive golf*. Chicago: Master Press.

Yu, B., Gabriel, D., Noble, L. and An, K. (1999). Estimate of the optimum cutoff frequency for the Butterworth low-pass digital filter. *Journal of Applied Biomechanics*, **15**, 318 – 329.

Zatsiorsky, V.M. and Aktov, A. (1990). Biomechanics of highly precise movements: the aiming process in air rifle shooting. *Journal of Biomechanics,* **23** (suppl. 1), 35-41.

Zhu, W. (1997). Making bootstrap statistical inferences: a tutorial. *Research Quarterly for Exercise and Sport,* **68**, 44-55.

APPENDICIES

Appendix A

Definition of measures used to indicate weight transfer

Fz% = Percentage of force under the front foot divided by the total force (Equation A1).

$$Fz\% = \frac{(100 * Fz2)}{(Fz1 + Fz2)} \qquad \text{Equation A1}$$

Where Fz1 = Fz under the back foot

Fz2 = Fz under the front foot

CP% (Centre of pressure) – Is the point at which the force vector intersects some defined plane, usually the ground surface. The calculations are represented below using CPy which is the direction usually associated with parallel to the line of shot (Equation A2).

$$CPy = \frac{Mx - (Fy * Dz)}{Fz} \qquad \text{Equation A2}$$

Where Mx = moment about the x-axis

Fy = force in the y-axis (horizontal)

Fz = force in the z-axis (vertical)

Dz = distance between transducer and a horizontal plane (usually the ground surface). For force plate 2 in this study (AMTI LG6-4): Dz2 = 0.0535 m.

COV (centre of vertical forces) – Equals the sum of the moments about a defined axis. For weight transfer parallel to the line of shot, the axis is perpendicular. For the Richards *et al*. (1985) study, a three transducer force plate was used (Figure A1). Examining Figure A1, moments will be calculated for each transducer, summed then divided by the total Fz to determine the horizontal coordinate for the force vector line of action.

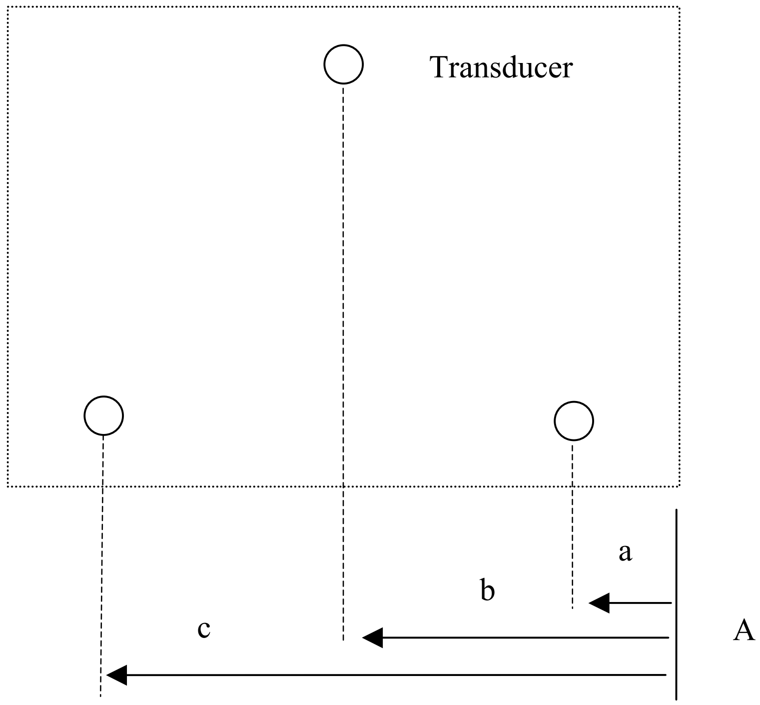


Figure A1: Calculation of COV using the sum of moments about A

Comparison of measures

Fz% and COV

The centre of vertical forces is equivalent to Fz% as it uses only vertical forces. The difference between the measures as used in the golfing literature is the Richards et al. (1985) study only used one force plate so the feet were required to be digitized so the resultant force vector position could be expressed relative to the feet. Fz% is simply calculated using equation A1 but could also be evaluated using equation A3 to obtain the same value.

Fz%/COV and CP

The difference between these measures is horizontal forces.

Expanding Equation A1

$$CPy = \frac{Mx - (Fy * Dz)}{Fz}$$

$$CPy = \frac{Mx}{Fz} - \frac{(Fy * Dz)}{Fz}$$

Fz% and COV equals the first part of this equation (i.e. Mx/Fz)
The contribution of the horizontal forces is the second part [(Fy*Dz)/Fz]

The difference between measures will depend on the magnitude of the horizontal and vertical force. However, the horizontal force is multiplied by Dz which is small (0.0535 m in the case of the AMTI LG6-4) reducing effect of Fy in the CP calculation. Using the mean Fy of 102N at BC (the largest Fy of any swing event) and the mean Fz for this study (Fz = 905 N), the change in CP value due to the horizontal force at this event is:

$$CPy = \frac{(100 * 0.0535)}{905} = 0.006m$$

So the difference between CP and the position of the vertical force Fz (i.e. Fz% or COV) will be less than 1 mm.

Appendix B

Comparison of Fz% and CPy%

The aim of this examination was to compare Fz% under each foot and CPy% between the feet. Both measures have been used to indicate weight position and transfer in golf studies.

Using the data from the sixty-two golfers in this study, CPy% between the feet and Fz% under each foot was quantified at eight swing events. The method of collecting CPy% data has been presented in the methods section (4.2.4.1). Fz% was calculated using equation B1.

$$\frac{Fz(UnderFrontFoot)}{Fz(UnderFrontFoot) + Fz(UnderBackFoot)} * 100 \qquad \text{Equation B1}$$

Correlations were performed to assess the relationship between Fz% and CPy% between the feet. Correlations were also performed between CPy% and Fz% data for each individual, with a mean correlation also obtained.

Observation of the group mean CPy% and Fz% data across the eight swing events indicated very similar patterns and values (table B.1 and figure B.1). This was supported by strong correlation coefficients between mean CPy% and Fz% on a group basis ($r = 0.999$, $p < 0.001$, $N = 62$). As well, weight position at the eight different swing events was similar, with a mean absolute difference of 1%. CPy% returned slightly lower values during backswing (MB, LB and TB) and slightly higher values during downswing (ED, MD and BC) and in follow through (MF).

Table B.1: Group means for CPy% and Fz% at eight swing events (*N* = 62)

	TA	MB	LB	TB	ED	MD	BC	MF
CPy%	57	28	22	23	63	70	71	66
Fz%	57	29	23	24	62	69	70	64

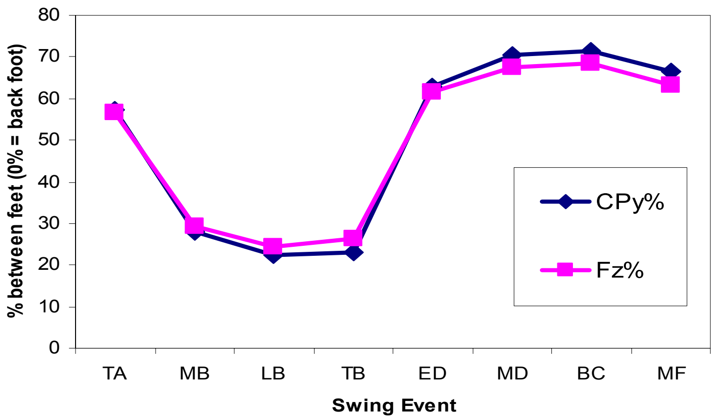


Figure B.1: Comparison of CPy% and Fz% between the feet during the golf swing (mean values at each event; *N* = 62 golfers).

CPy% and Fz% measures also showed strong similarities on an individual basis. The mean correlation for all individuals between CPy% and Fz% was $r = 0.996$ (range $r = 1.000 – 0.947$; all significant at $p < 0.001$). There were five of golfers who showed large differences (> 5%) at certain events, particularly at TB and MF where horizontal forces were higher.

As the correlations between CPy% and Fz% were strong for this group, statistical analyses involving either measure would be expected to be similar. This means that statistical analyses for studies using either measure can be compared with confidence that the type of measure is not influencing the data. For example the correlations between Club Velocity and CPy% range ($r = 0.11$, $p = 0.39$) and Club Velocity and Fz% range ($r = 0.12$, $p = 0.36$) were very similar. The only time this might have been an issue was where the measures were close to a threshold value (e.g. 0.2 as a cut-off for small effect compared with no effect). However, this did not occur in any measure for this study.

As mean differences between CPy% and Fz% values were low (≤ 3% on a group basis) comparison of CPy% and Fz% values may also be performed when examining group-based data. However, as some individuals produced large (up to 5%) differences, the comparisons on an individual basis are likely to hold more error and should be treated with caution.

Appendix C

Smoothing

Observation of the raw force plate data for golf swings and for static loading situations (weights placed on force plate) indicated a high frequency noise existed in force and moment data that required smoothing. Figure C.1 shows a spectral analysis of CPy when the force plates were loaded with 750 N of weight, showing a relatively large 50 Hz spike with low amplitude noise across the frequency spectrum which was slightly larger amplitude between approximately 30 Hz and 60 Hz.

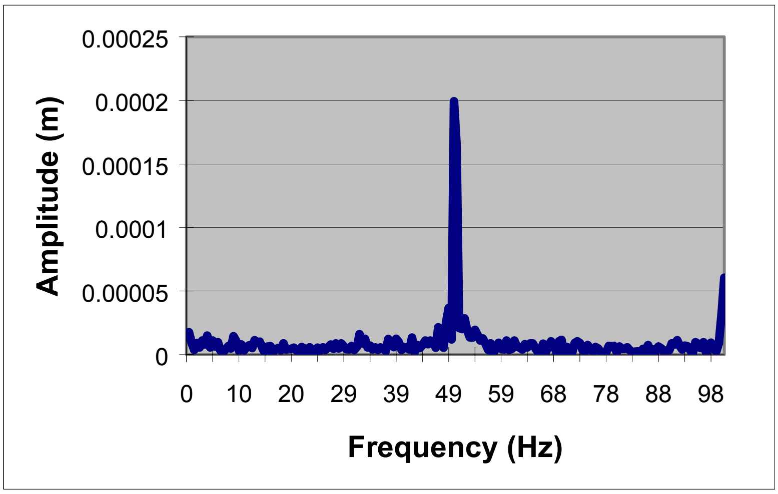


Figure C.1: Spectral analysis of CP with a 750 N weight placed on the force plate
Note: Calculation on 1024 samples (2.048 s at 500 Hz). Unusual time interval due to the FFT calculation requiring a value which is a power of 2.

Based on these observations, a low pass pre-filter (24.3 Hz) was inserted into the AMLAB software with the aim to eliminate this noise. While no frequency domain data has been presented in the literature for weight transfer in the golf swing, it was considered that no frequencies above this level would be expected in weight transfer data in the golf swing.

However, noise still existed in the system, evident in CPy displacement curves in figure C.2 calculated from data sampled with the 24.3 Hz pre-filter. Spectral analysis indicated that some of this noise was 50 Hz, indicating that noise was added after the pre-filtering but before the data was stored. This researcher assumed the source of this noise was the analogue to digital computer board. As such, it was decided that smoothing the CP displacement data was required.

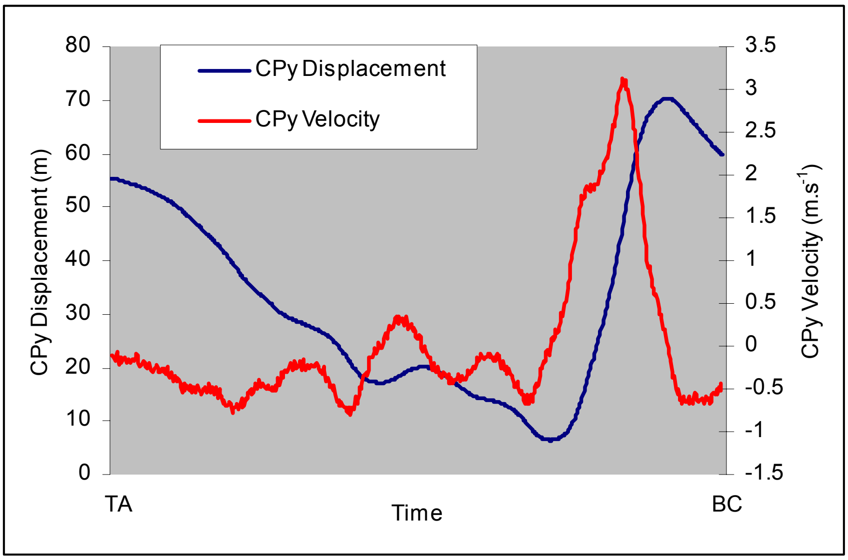


Figure C.2: CP displacement during a golf swing calculated from force plate data pre filtered at 24.3 Hz.

CP Displacement

To decide on an appropriate smoothing cut-off frequency for CP displacement and velocity, a combination of methods was used to gather information on the data. These were:

1. Automatic methods for determining optimal smoothing frequency.
2. The effect of different smoothing cut-off frequencies on parameters of interest.
3. Observation of raw and smoothed data curves (displacement and speed)

Ball et al. (2001) recommended a combination of the above methods as well as spectral analysis for thorough assessment of smoothing requirements. However, as the golf swing is non-stationary, accurate spectral analysis would have required the use of wavelets. Wavelet software was unavailable to this researcher at the time of deciding upon a smoothing cut-off frequency and development of this software was considered beyond the scope of this study. As well, no frequency domain data for the golf swing exists in the literature so the decision of smoothing cut-off frequency was made based on methods 1-3 only.

1. Automatic methods for determining optimal smoothing frequency

Three automated methods for calculating an optimal smoothing cut-off frequency were applied to CP data; Challis (1999), Yu *et al.*, (1999) and Winter (1990). Table C.1 reports the mean cut-off frequencies returned by each of the methods (5 golf swing trials from randomly selected golfers examined). CPx has been included as it was evaluated prior to analysis of the data and was used in comparison with Neal (1998) in study 1.

Table C.1: Cut-off frequencies found by different methods to be optimal.

	Displacement	
	CPx	CPy
Challis (1999)	14.0	15.2
Yu et al. (1999)	24.8	24.9
Winter (1990)	14.0	15.0

The Challis (1999) method and the Winter (1990) method produced similar cut-offs for both CPy and CPx. The Yu et al. (1999) method returned larger values than the Challis (1999) and Winter (1990) methods of approximately 25 Hz, due to the large sample rate in this study (500Hz; The Yu method is largely sample rate based).

2. The effect of different smoothing cut-off frequencies on parameters of interest.

Inspection of parameters of interest (CPy% at swing events, maximums and minimums) indicated that for most parameters, minimal change existed using cut-off frequencies from 15-25 Hz. At 10 Hz, some change was evident and at 5 Hz the change was large for some parameters. Figure C.3 shows an example of the effect of different smoothing cut-off frequencies on the values of CPy% maximum and CPy% minimum for a single trial. At 10 Hz, CPy% maximum began to change more considerably than for the higher frequencies, indicating that the smoothing is influencing values and that signal as well as noise may be being eliminated. While the differences in these parameters are relatively small (Change in CPy% < 1%), for other parameters (e.g. velocity) this change was more considerable although similar patterns existed where below 15 Hz, parameter values changed.

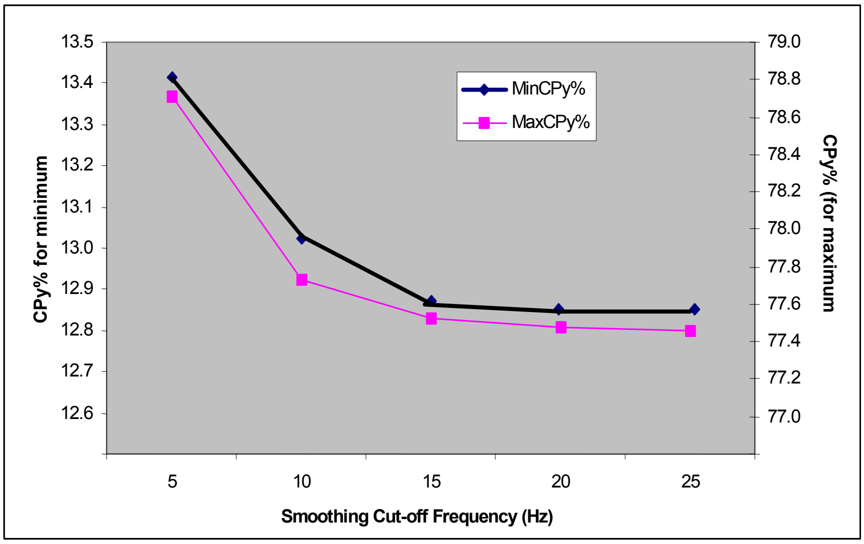


Figure C.3: Effects of different cut-off frequencies on parameters of interest (CPy% maximum and CPy% minimum)

3. Observation of raw and smoothed data curves (displacement and velocity)

Figures C.4 and C.5 show raw and smoothed CPy displacement and CPy velocity curves for a single trial across the whole swing from address to mid follow through. The same data is presented again for downswing only (TB – MF) to enable better inspection of the changes due to smoothing. Inspection of displacement curves smoothed at a range of frequencies from 5 Hz to 30 Hz indicated that little change existed between raw and smoothed data from 15-30 Hz. Slight changes were noted on the later stages of downswing at 5 Hz and 10 Hz, which were more particularly noticeable in velocity data at the peaks. This supported the findings of parameter changes from the previous section.

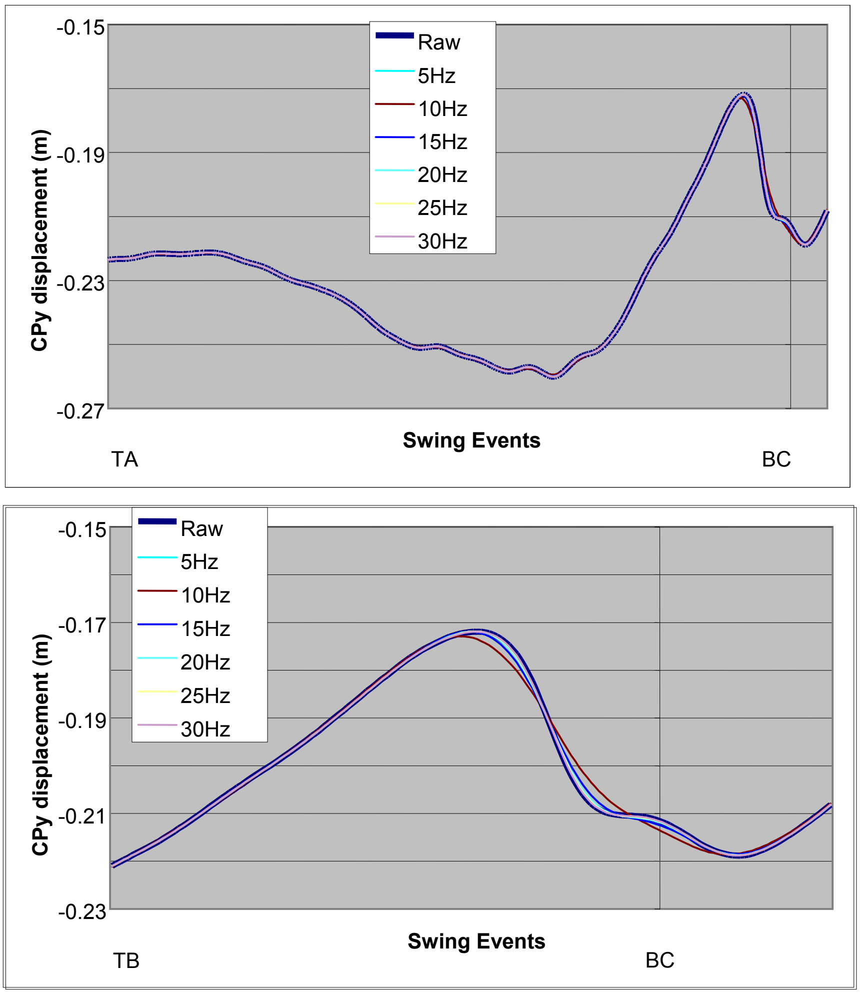


Figure C.4: Example CP displacement data: raw and smoothed at different cut-off frequencies (same curve with second graph focusing on downswing).

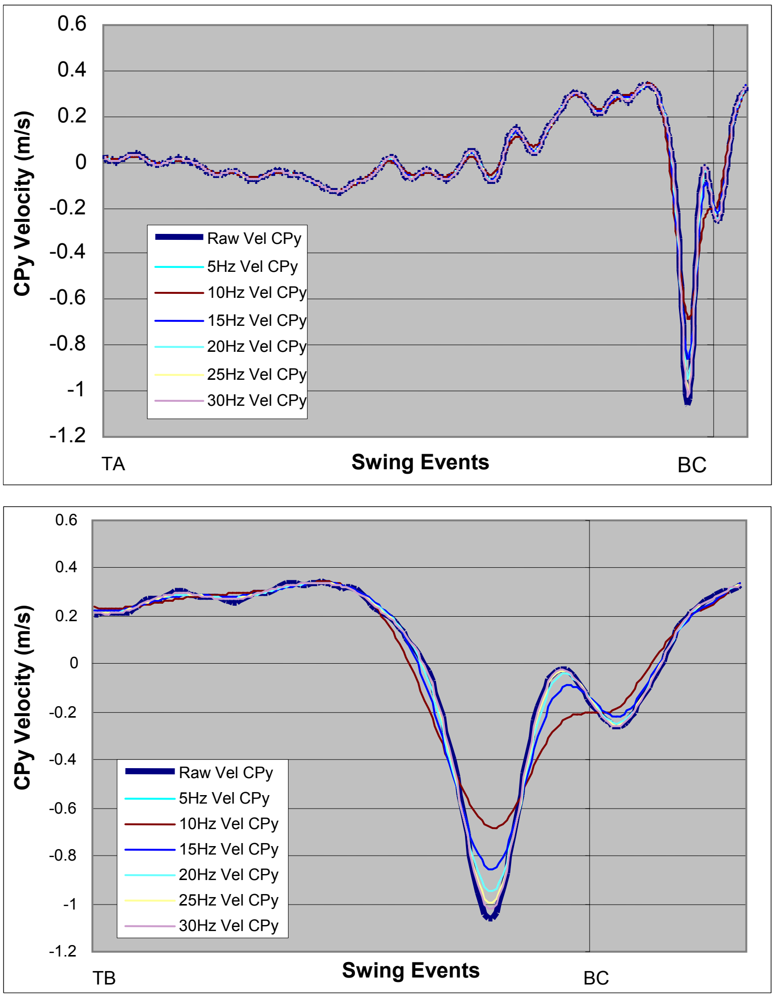


Figure C.5: Example CP velocity data: raw and smoothed at different cut-off frequencies (same curve with second graph focusing on downswing).

Smoothing summary

It was decided to use a 15 Hz cut-off for both displacement and velocity data for this study. This cut-off was chosen based on three levels of decision making as recommended by Ball *et al*. (2001). First, 15 Hz was indicated as optimal by two of three automated algorithms (Challis, 1999; Winter, 1990; Yu *et al*., 1999). The Challis (1999) and Winter (1990) methods produced similar cut-offs for CPy of 15.2 and 15.0 respectively. The Yu *et al*. (1999) method returned substantially larger values of approximately 25 Hz, due to the large sample rate in this study (500Hz; The Yu *et al*. method is strongly influenced by sample-rate). Second, the influence of different cut-offs on parameters of interest (CPy% between the feet and CPy velocity at swing events as well as maxima and minima) was inspected. Large changes in parameter values were evident when cut-offs below 10 Hz to 15 Hz were used. This level was considered to represent oversmoothing. Third, visual inspection of raw and smoothed curves indicated the 15 Hz cut-off provided smooth displacement and velocity curves without attenuating what was considered real data in particular near the maxima and minima.

Appendix D

Validation of ProV system

To validate the ProV system, Club Velocity obtained by the ProV was compared with Club Velocity calculated from digitised data.

Twenty swings were performed with a 5-iron. For each swing, Club Velocity was recorded by the ProV system. Also for each swing, video from an overhead 200 Hz video camera, aligned perpendicular to the hitting surface and immediately to the side of the point light source for the ProV was obtained (figure D.1). The two frames immediately before ball contact were used for analysis using Peak Motus. For each trial, two points on the clubface were digitised; one near the heel and one near the toe of the club. The average of the two points was calculated for each frame and velocity of this average point was calculated and used to indicate Club Velocity.

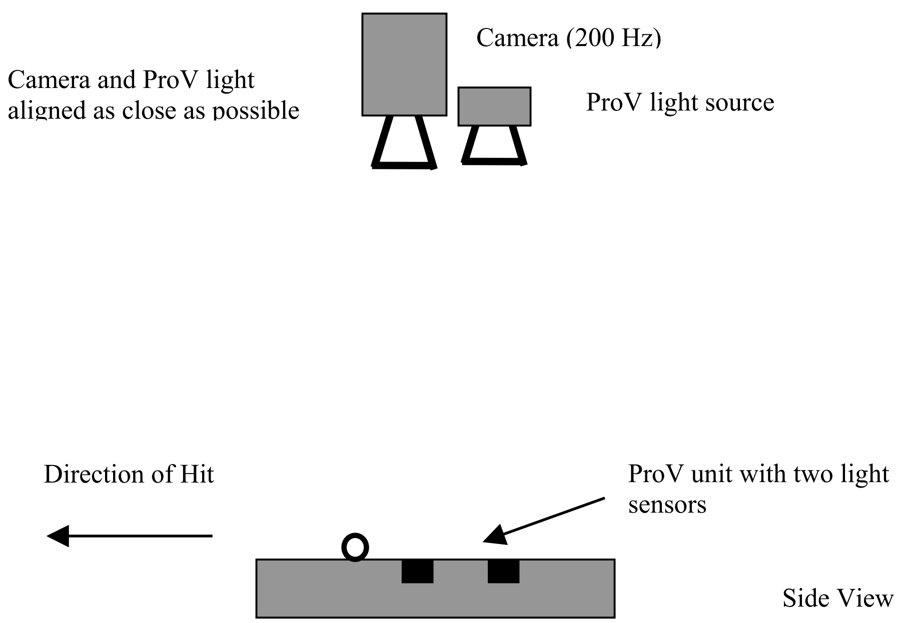


Figure D1: Laboratory set-up for comparison of Club Velocity

Table D.1 reports the ProV and digitised data for club velocity in both feet per second (units in which ProV specified error) and metres per second (used in this study) have been reported.

Table D.1: Comparison of Club Velocity obtained from digitising and from the ProV system

	ProV Data		Digitised Data		Difference	
	$ft.s^{-1}$	$m.s^{-1}$	$ft.s^{-1}$	$m.s^{-1}$	$ft.s^{-1}$	$m.s^{-1}$
	95.9	29.3	96.3	29.4	-0.4	-0.1

Club Velocity (N = 20)

The mean difference in Club Velocity between the digitised and ProV data (0.4 $ft.s^{-1}$) was within the factory specified error of +/- 0.5 $ft.s^{-1}$. As well, although not reported here as the data has not been used in this study, differences between ProV and digitised data for clubface angle and swing path, two other measures produced by ProV, were also within the factory specified error. The difference between ProV data and digitised data was not significant (F = 0.62, p = 0.78).

It should be noted that the method of validation used here is not a gold standard comparison as such. Rather this is a comparison of an accepted measure of velocity (from digitised data) with that of the ProV. The use of only two frames, rather than a series of frames which was then smoothed, was the preferred method for three reasons. First, it allowed for measurement at the same site (i.e. over the appropriate ProV lights). Second, it allowed for a larger image for digitising as the field of view could be narrowed to the area immediately near the ball before contact. If more frames were to be digitised, the field of view would have been required to be increased, reducing the resolution of the digitising set-up (i.e. 1 pixel would have equalled a greater distance so a 1 pixel error would be greater in the video with the wider field of view). Third, the 200 Hz camera did not capture enough points for a reasonable extrapolation across which smoothing could be performed. Even when the field of view was increased to its maximum, the golf club was in view for only 8 – 10 video fields. As there is an impact in the movement, the data needed to be examined in two sections – pre and post impact (Knudson and Bahamonde, 2001). To perform this analysis, the most common method in the literature is to use data up to the instant before impact then extrapolate from this point to add more data points prior to smoothing. However, due to the rapid movement and relatively slow frame rate (at 200 Hz, the club was only evident over the ProV system for 2 - 4 frames) the use of this method was not appropriate.

In summary, ProV Club Velocity measurement lay within the ProV quoted error limits of 1 $ft.s^{-1}$. Comparison with digitised velocity data produced non-significant differences of only 0.4 $ft.s^{-1}$. It was concluded that the data provided by the ProV system was precise enough to be used in this study.

Appendix E

Summary of error assessment

Table E.1 (200 Hz camera) and table E.2 (250 Hz camera) present the error associated with each parameter used in this study. It is presented here to allow easy reference for the reader. Calculations for error estimation for each parameter are provided in the next sections. In all cases, means for significantly different parameters were well in excess of these errors. Further, cluster, correlation and regression analysis were not influenced by these errors as values were rounded to the appropriate error level prior to analysis, although mean data has been reported to a slightly higher level in some cases for easier reading. For example it would have been confusing to report values in units of 0.3 % for CPy% rather than in whole units of 1% (or in this case, 0.1%)

Table E.1: Approximate and maximum error estimates for parameters used in this study (200 Hz camera). All values +/-.

Performance	Approximate Single Measure	Approximate Across 10 trials	Maximum
Club Velocity ($m.s^{-1}$)	0.1	0.04	0.1
At each swing event (% between the feet)			
CPy%TA	0.6	0.2	2.1
CPy%MB	0.7	0.2	2.6
CPy%LB	0.6	0.2	2.3
CPy%TB	0.7	0.2	3.5
CPy%ED	1.0	0.3	3.8
CPy%MD	0.7	0.2	4.0
CPy%BC	0.6	0.2	1.7
CPy%MF	1.0	0.3	4.6
Average	0.7	0.2	3.0
CPy Velocity ($m.s^{-1}$)			
At downswing events			
VelCPyTA	0.18	0.05	7.00
VelCPyMB	0.18	0.04	7.00
VelCPyLB	0.18	0.04	7.00
VelCPyTB	0.18	0.06	7.00
VelCPyED	0.23	0.15	7.00
VelCPyMD	0.20	0.10	7.00
VelCPyBC	0.18	0.04	7.00
VelCPyMF	0.23	0.16	7.00
Average	0.19	0.08	7.00
VMaxCPy ($m.s^{-1}$)	0.18	0.06	0.25
tVMaxCPy (s)	0.001	< 0.001	0.002
MaxCPy% (% between the feet)	0.56	0.18	1.6
tMaxCPy% (s)	0.001	< 0.001	0.002
MinCPy% (% between the feet)	0.56	0.18	1.6
tMinCPy% (s)	0.001	< 0.001	0.002
CPyR (m)	0.007	0.002	0.01
CPyR% (% between the feet)	0.79	0.25	1.12

Table E.2: Approximate and maximum error estimates for parameters used in this study (250 Hz camera). All values +/-.

Performance	Approximate Single Measure	Approximate Across 10 trials	Maximum
Club Velocity ($m.s^{-1}$)	0.5	0.16	0.5
At each swing event (% between the feet)			
CPy%TA	0.6	0.2	2.0
CPy%MB	0.6	0.2	2.4
CPy%LB	0.6	0.2	2.1
CPy%TB	0.6	0.2	3.1
CPy%ED	0.8	0.2	3.3
CPy%MD	0.7	0.2	3.5
CPy%BC	0.6	0.2	2.2
CPy%MF	0.8	0.3	3.9
Average	0.6	0.2	2.8
CPy Velocity ($m.s^{-1}$)			
At downswing events			
VelCPyTA	0.18	0.05	7.00
VelCPyMB	0.18	0.04	7.00
VelCPyLB	0.18	0.04	7.00
VelCPyTB	0.18	0.05	7.00
VelCPyED	0.21	0.12	7.00
VelCPyMD	0.19	0.09	7.00
VelCPyBC	0.18	0.06	7.00
VelCPyMF	0.21	0.13	7.00
Average	0.19	0.07	7.00
VMaxCPy ($m.s^{-1}$)	0.18	0.06	0.25
tVMaxCPy (s)	0.001	< 0.001	0.002
MaxCPy% (% between the feet)	0.56	0.18	1.6
tMaxCPy% (s)	0.001	< 0.001	0.002
MinCPy% (% between the feet)	0.56	0.18	1.6
tMinCPy% (s)	0.001	< 0.001	0.002
CPyR (m)	0.007	0.002	0.01
CPyR% (% between the feet)	0.79	0.25	1.12

Approximation of error in the parameters used in this study was difficult due to the need to combine data obtained from three different measurement systems (force plate data, digitized data and timing data) and the lack of a gold standard for comparison. A combination of experimental and theoretical methods was used to determine an approximate error for each parameter.

Techniques presented by Taylor (1982) for calculating measurement error that is propagated during calculations have been used. Briefly, these are:

- *Uncertainty in sums and differences*: if two terms with error in them are added or subtracted, the error in each term is added
- *Uncertainty in a measured quantity multiplied or divided by an exact number*: if a term is multiplied or divided by an exact number, the error in the term is multiplied or divided by the constant
- *Uncertainty in products or quotients*: if two terms with error in them are multiplied or divided, the error is calculated by using fractional uncertainties.

For example, to calculate the error in a quotient, the error is calculated as a fraction of the measured value for the upper line and the lower line separately, adding them to get a total fractional uncertainty and then calculating the error value using the final measured value (this is more easily understood in the calculations below – see error calculations for CPy% between the feet)

- Quadrature summation: If a value is entered more than once in a calculation or comes from the same source (e.g. two CPy% values used to calculate CPyR%), some cancelling of error can be expected. Taylor reported a better way of approximating error to take into account this cancelling effect called quadrature summation. This involves squaring the error terms, summing them and then taking the square root of this sum. Approximate error has been estimated using this process.

Note: the symbol 'δ' is used to indicate 'error in..' and is used before terms to denote that the error (rather than the value itself) is being referred to.

Force Plate error analysis

In order to assess how accurately the force plate system located the CP, a comparison between force plate calculated CPy coordinates was made with known CPy coordinates.

A grid of masking tape was secured to each force plate and a point was marked with a pen at the middle of each intersection. These grids provided 8 points on the OR6-5-1 force plate and 18 points on the LG6-4 force plate and were spaced out evenly across the force plates (see table E.3 and E.4 for coordinates). Each position was measured using a millimetre ruler to obtain the y coordinates. A javelin was used to provide a point source of force at each of the grid positions. The tip of the javelin was pressed into the force plate at each grid marking with a vertical force of between 200 N and 230 N, similar to the force values used by Sommer *et al.* (1997; similar methodology to evaluate CP in Kistler force plates). CP was calculated for each position in Excel and data was compared with the known coordinates. Similarity between the two measures was used to indicate accuracy. That is, CP positions recorded by the force plate system should be the same as the grid coordinates.

Table E.3: Large force plate (LG6-4) grid data showing force plate CP and grid coordinates difference (all data in mm)

Grid Position	Force Plate Data (CP)	Grid Coordinates	Absolute Difference
1	1047.1	1048.5	1.4
2	1047.9	1049.5	1.6
3	1047.1	1050.5	3.4
4	898.3	900.0	1.7
5	899.5	899.5	0.0
6	899.1	899.0	0.1
7	749.7	749.0	0.7
8	749.8	749.0	0.8
9	750.4	750.0	0.4
10	599.4	600.0	0.6
11	599.8	600.0	0.2
12	598.3	599.5	1.2
13	449.8	449.5	0.3
14	449.2	450.0	0.8
15	448.4	449.0	0.6
16	301.0	300.0	1.0
17	300.2	299.0	1.2
18	299.1	299.5	0.4
Mean Error			**0.9**

Table E.4: Small force plate (OR6-5-1) grid data showing force plate CP and grid coordinates difference (all data in mm)

Grid Position	Force Plate Data (CP)	Grid Coordinates	Absolute Error
1	398.0	399.0	1.0
2	299.9	300.0	0.1
3	200.8	201.5	0.7
4	100.1	100.0	0.1
5	101.7	100.5	1.2
6	199.1	200.5	1.4
7	298.1	301.0	2.9
8	401.3	401.0	0.3
Mean Error			**1.0**

Average difference between measures for both force plates were small (approximately 1.0 mm) and within the precision available to the analysis (i.e. javelin tip approximately 1 mm wide, grid coordinates measured to ±0.5 mm). A few areas of the large force plate returned large differences (e.g. Position 3, error = 3.4 mm). These were near the edge of the force plate and are similar to the error patterns reported by Sommer *et al.* (1997) in Kistler force plates. This prompted the testers to position golfers away from the edges and towards the middle of the force plates.

This error is smaller than the error reported for Kistler force plates (e.g. Sommer *et al.*, 1997; Bobbert and Schamhardt, 1990; Middleton *et al.*, 2000: maximum possible error of 20 mm; mean error approximately 5 mm). This may be due to the plate distortion characteristics of the AMTI force plates, which use strain gauge technology (which require distortion for measurement), affecting the measurement less than the Kistler force plates, which use Piezo electric crystals with a solid metal top plate and require the two surfaces (top plate and piezoelectric crystal) to remain perpendicular. It might also be due to the influence of a point source load, as opposed to an area load that will most likely be the source of force applied to the force plate (e.g. the sole of a shoe). Schmeidmayer and Kastner (2000) and Middleton *et al.* (1999) found that Kistler force plates did not show as large an error in CP measurement if the load applied was in the form of an area rather than a point source, as used by Bobbert and Schamhardt (1990) and Sommer *et al.* (1997), suggesting that the force plate will distort differently and as such, the error will be different. For example Middleton *et al.* report finding errors of approximately 5 mm using a point source which were reduced to less than 2 mm when two metal blocks were placed on the force plate, similar to the area loading that might be expected when a subject is standing on the force plate. Regardless, using similar point force methods, the AMTI force plate data produced smaller errors that Kistler force plates for CP. No experimental data exists for AMTI force plates in the literature.

Chockalingham *et al.* (2002) reported errors greater than 3 mm below 90N in single force plate analysis. While force values might get low on the back foot near ball contact, the total Fz remains large (between 0.95 and 1.3 times body weight) for the entire golf swing (as opposed to the gait problem examined by Chockalingham *et al.*). As such, the relative effect of any error is reduced substantially in this study. This is because the CP calculation included information on Fz from both plates (see equation below) and if Fz approaches zero on one plate it is multiplied by the erroneous CP value and the sum (eg. Fz2*CPy2 – see equation below) approaches zero.

$$CPy = \frac{(Fz1 * CPy1) + [Fz2 * (CPy2 + Df2)]}{Fz1 + Fz2}$$

Note that an extensive error assessment of the force plate system has been conducted examining CP under different force conditions. As this was presented in Ball (1999, unpublished Masters thesis, Victoria University) it is not appropriate to present it here as new work.

In summary, the mean error for each force plate system was used to indicate error in CP measurement (LG6-4: 0.9 mm or +/- 0.45 m; OR6-5-1: 1.0 mm or +/- 0.5 mm).

Foot Position

This process has been reported elsewhere in detail (Brown, 2002) and results are summarized in table E.5. Briefly a golfer was set-up in the address position on the hitting area as for testing. The heel and toe position of each foot were measured manually using a ruler (graded in mm). The image was recorded for analysis. This error was considered the error in a single digitized field. As four fields were digitized, error will be reduced due to cancelling of the random error associated with the digitizing process. The approximate error, calculated using quadrature summation, is also included in this table (calculations shown after the table).

Table E.5: Maximum and approximate error for foot position (+/- mm)

	Back Foot Toe	Back Foot Heel	Front Foot Toe	Front Foot Heel
Maximum error	4	1	5	6
Approximate error	2.0	0.5	2.5	3.0

Approximate error calculation

Using the Front Foot heel as an example (all values calculated using the same method):

$$CPy_{heel} = \frac{y_{heel}A + y_{heel}B + y_{heel}C + y_{heel}D}{4}$$

Where A, B, C and D represent the four digitised fields.

Approximate error in Front Foot heel position

$$= \frac{\sqrt{y_{heel}A^2 + y_{heel}B^2 + y_{heel}C^2 + y_{heel}D^2}}{4}$$

$$= \frac{\sqrt{6^2 + 6^2 + 6^2 + 6^2}}{4} \qquad = 3 \text{ mm}$$

Note: Maximum possible error will remain the same as in table E.5 (i.e. each measure will be in maximum error in the same direction)

Error in the Front Foot position

Front Foot position $= \dfrac{y_{heel} + y_{toe}}{2}$

Maximum error in Front Foot position $= \dfrac{\delta y_{heel} + \delta y_{toe}}{2}$

$$= \frac{6+5}{2} \qquad = 5.5 \text{ mm}$$

Approximate error in Front Foot position $= \dfrac{\sqrt{\delta y_{heel}{}^2 + \delta y_{toe}{}^2}}{2}$

$= \dfrac{\sqrt{3^2 + 2.5^2}}{2}$ = 2.0 mm

Error in the Back Foot position

Back Foot position $= \dfrac{y_{heel} + y_{toe}}{2}$

Maximum error in Back Foot position $= \dfrac{\delta y_{heel} + \delta y_{toe}}{2}$

$= \dfrac{1+4}{2}$ = 2.5 mm

Approximate error in Back Foot position $= \dfrac{\sqrt{\delta y_{heel}{}^2 + \delta y_{toe}{}^2}}{2}$

$= \dfrac{\sqrt{2^2 + 0.5^2}}{2}$ = 1.0 mm

Error in CPy%

CPy% between the feet =

$$\frac{CPy - BackFootPosition}{FrontFootPosition - BackFootPosition}$$

Using the error calculated for the back and front foot positions and recalling the error in CPy was measured as +/-0.5 mm and separating the equation to calculate error:

Upper line $= CPy - BackFootPosition$

Error in upper line $= \delta CPy + \delta BackFootPosition$

Maximum = 0.5 + 2.5 = 3.0 mm

Approximate = 0.5 + 1.0 = 1.5 mm

Lower line $= FrontFootPosition - BackFootPosition$

Error in lower line $= \delta FrontFootPosition + \delta BackFootPosition$

Maximum = 5.5 + 2.5 = 8.0 mm

Approximate = 2.0 + 1.0 = 3.0 mm

Overall error in CPy% between the feet

For this calculation, measured data was required (e.g. position of feet, displacement between feet). Five trials from different golfers were obtained and all CP parameters were calculated using the same procedures as in study 1. The mean of the five trials for each parameter was used in error calculations. These are reported as they become necessary in the calculations. Percentage errors are used in calculation of error for each parameter. This should not be confused with the reported error for the parameters that are expressed as a percentage. For example, CPy% error is reported as 0.6%. This is the value of the error (i.e. the error is not 0.6% of the measurement).

Mean position of back foot = 205 mm
Mean position of front foot = 706 mm
Let CPy = 605 mm

CPy% between the feet $= \dfrac{CPy - BackFootPosition}{FrontFootPosition - BackFootPosition} * 100$

$= \dfrac{605-205}{706-205} * 100 \qquad = 80\%$

δCPy% between the feet $= \left|\dfrac{\delta UpperLine}{UpperLineValue}\right| + \left|\dfrac{\delta LowerLine}{LowerLineValue}\right|$

Maximum (% of CPy%) $= \left|\dfrac{3}{605-205}\right| + \left|\dfrac{8}{706-205}\right| \qquad = 0.023$

Maximum error = CPy% * maximum (% of CPy%)
= 80 * 0.023 = 1.6 %

Approximate (% of CPy%) $= \left|\dfrac{1.5}{605-205}\right| + \left|\dfrac{3}{706-205}\right| \qquad = 0.009$

Approximate error = CPy% * approximate (% of CPy%)
= 80 * 0.009 = 0.6 %

So a single measure of CPy% can be in error by +/- 0.6 % with a maximum possible error of 1.6 %. The approximate error was considered reasonable for this study and while the maximum possible error was large, it is unlikely that this value would exist due to the cancelling effects throughout the calculations and the likelihood that digitizing errors would be reduced by the use of four fields for analysis.

Timing Errors

This analysis was performed by another researcher in this study and will be presented fully in that thesis. This section provides a brief outline of this process.

Error in obtaining timing data could have occurred in two areas

1. Error in identification of the correct video field was evaluated.
2. Error in the precision supplied by the 200 Hz and 250 Hz cameras

1. Error in identification of the correct video field was evaluated.

For ten randomly selected trials, full digitization on Peak MOTUS was performed. Two points on the shaft of the club were identified (reflective tape was applied to the club shaft near the handle and near the club head). The field nearest each of the eight swing events was identified from this data as per study 1 (table 4.2.4.2.1, section 4.2.4.2). For example, using the screen coordinate system, the field where the club shaft was nearest to horizontal in the downswing was considered the field indicating mid downswing (MD). In the case of the top of backswing, the field before the first change in angular velocity direction of the club was identified.

The tester then identified the eight swing events subjectively from the video image, as was performed for the study.

The error in identification was the difference between the two methods, with the digitized method considered the 'gold standard'.

There was no difference between the digitized data and subjectively identified data for all events for all trials with the exception of TA and TB. These events were found to be in error by +/- 1 video field (or +/- 0.005 s for the 200 Hz camera and +/- 0.004 s for the 250 Hz camera).

2. Error in the precision supplied by the 200 Hz and 250 Hz cameras

As identification of most swing events was exact, the error associated with the use of swing events lay only in its precision. At 200 Hz, the exact event did not always occur at the same time as a video field was captured. For example, one video field might have been captured just before ball contact while the next captured just after ball contact, with the event itself occurring between fields. As such, the error in CPy% could be up to half the difference in CPy% between video fields. Due to the video and CPy% sampling rates being different, there were 2.5 CPy% data points for every 1 video field. So the error due to limits of the video could be +/- 1.25 CPy% data points. This error in field identification could occur at the event itself or at ball contact (as the time was synchronized to ball contact). Maximum error would occur if the swing event and ball contact lay exactly between fields and the earlier field was chosen for the swing event and the later field was chosen for ball contact (or vice versa). In this case, the total error would be equivalent to 1 field or 2.5 CPy data points (2 sample points for the 250Hz camera). Maximum error for TA and TB was the sum of this error plus the error in identifying a field either side of the true field (or +/- another 2.5 CPy data points for a total of +/- 5 CPy data points).

To calculate the error due to timing, five trials from five different golfers were chosen at random. For each golfer, CPy% and CP velocity was calculated for the entire swing. At each swing event, the measured value was obtained (i.e. the value indicated as at that swing event) as well as 3 CPy% (and CPy velocity) data points either side of this data point (2 samples for the 250Hz camera). The difference between the measured value and the value at +/-3 samples was averaged for each golfer and used as indicative of error due to timing. Note that the use of 3 samples is a slight overestimate of the error (1 field = 2.5 samples for the 200 Hz camera) but the 'true' error lay between samples and may have been two or three samples at different times, depending on the exact point of ball contact). It should be noted that where a possible error lay at a fraction of a field, the larger error was chosen (e.g. a 2.5 data point error would be examined as a 3 data point error)

Table E.6 and E.7 reports the mean and single maximum error for each swing event from five randomly selected golfers using the 200Hz and 250 Hz cameras.

Table E.6: Mean and single maximum error for each swing event from five randomly selected golfers (200 Hz camera). All values +/-.

	Mean Error		Maximum Error	
Swing Event	CPy%	Velocity ($m.s^{-1}$)	CPy%	Velocity ($m.s^{-1}$)
TA	0.23	0.02	0.29	0.04
MB	0.35	0.01	0.58	0.02
LB	0.15	0.01	0.41	0.01
TB	0.37	0.03	1.09	0.05
ED	0.83	0.08	1.27	0.16
MD	0.49	0.06	1.40	0.10
BC	0.05	0.01	0.14	0.02
MF	0.90	0.09	1.04	0.11
AVERAGE	0.42	0.04	0.83	0.06

Table E.7: Mean and single maximum error for each swing event from five randomly selected golfers (250 Hz camera). All values +/-.

	Mean Error		Maximum Error	
Swing Event	CPy%	Velocity ($m.s^{-1}$)	CPy%	Velocity ($m.s^{-1}$)
TA	0.14	0.01	0.17	0.03
MB	0.21	0.01	0.35	0.01
LB	0.09	0.00	0.25	0.01
TB	0.22	0.02	0.65	0.03
ED	0.50	0.05	0.76	0.10
MD	0.49	0.03	0.84	0.06
BC	0.08	0.01	0.21	0.02
MF	0.58	0.05	1.00	0.06
AVERAGE	0.29	0.02	0.52	0.04

As can be noted in these tables, different errors occurred at different swing events. This was due to the different rates of change of CPy% and CPy velocity. The largest errors occurred at ED, MD and MF, stages of the swing where the club and body is moving most rapidly. BC error is low due to the force plate system being triggered at BC and identification from video was not used in its calculation (hence error is only +/- 0.5 data points).

It should be noted for these calculations that the error is probably smaller than the values reported, in most events, the estimation of when the swing event occurred could be determined 'between fields' and so providing a 400 Hz sample rate. For example, it was relatively easy to identify if ball contact occurred closer to a particular field or closer to the midpoint between fields. This was assessed against digitized data also and was found to be appropriate for downswing events (but not backswing events). However, it was felt that it was more appropriate to determine the error based on the sample rate provided by the system but with the knowledge that this represents an upper limit of error.

Overall Error

CPy% at eight swing events

Table E.8 and E.9 report the error in CPy% at each swing event using the 200 Hz and 250 Hz cameras. Error was calculated as:

Error due to foot position/force plate data (CPy%) + error due to timing (at each event)

Approximate error calculated using quadrature summation while maximum error calculated as the sum of both errors.
Recalling
Error due to force position/force plate data (CPy%) = +/- 0.6% (maximum = 1.6%)
Error due to timing from data in tables E.6 and E.7 (tables from previous section).

Table E.8: Error in CPy% at each swing event (200 Hz camera). All values +/-%.

	Approximate		Maximum
	Single measure	Over 10 measures	Single maximum error
TA	0.6	0.2	2.1
MB	0.7	0.2	2.6
LB	0.6	0.2	2.3
TB	0.7	0.2	3.5
ED	1.0	0.3	3.8
MD	0.7	0.2	4.0
BC	0.6	0.2	1.7
MF	1.1	0.3	4.6
AVERAGE	0.7	0.2	3.1

Table E.9: Error in CPy% at each swing event (250 Hz camera). All values +/-%.

	Approximate		Maximum
	Single measure	Over 10 measures	Single maximum error
TA	0.6	0.2	2.0
MB	0.6	0.2	2.4
LB	0.6	0.2	2.1
TB	0.6	0.2	3.1
ED	0.8	0.2	3.3
MD	0.7	0.2	3.5
BC	0.6	0.2	2.2
MF	0.8	0.3	3.9
AVERAGE	0.6	0.2	2.8

A further calculation is also presented in tables E.8 and E.9. As the CPy% value for each golfer was averaged over 10 trials, the random error in the value used for each golfer in study 1 and study 2 will be lower than that reported for a single swing. The error value over 10 measures was calculated using quadrature summation of the single measure error:

Approximate error across 10 trials $= \dfrac{\sqrt{\delta one.trial^2 * No.oftrials}}{No.oftrials}$

$$= \frac{\sqrt{0.6^2 * 10}}{10} = 0.2\%$$

The maximum remains the same regardless of how many trials are performed and would be achieved if every trial were in error by the maximum amount and in the same direction (in which case it would be a systematic and not a random error).

As mentioned, although the maximum values are large in some cases (e.g. MD = 4%) it is very unlikely that this error will occur due to canceling effects of errors throughout calculations and the averaging process. Further, in study 1 and study 2, group means were associated with N of greater than 15 for the large clusters. Group mean error for fifteen golfers or more would be approximately 0.05% for CPy%. As such it as decided to report these means to 0.1% increments. Post hoc evaluation indicated that all significantly different groups produced means well in excess of approximate and maximum errors.

CP velocity

CP velocity presented a problem in that there was no method for producing a known velocity to compare against. While the value of +/-0.5 mm was used for CPy in the calculation of error it probably doesn't offer a reasonable estimate as data is smoothed twice between measurement of CP and the production of velocity. As such, the use of an approximate displacement resolution has limitations. A second method (theoretical) was also performed using the resolution of the 16-bit ADC system to define the error in CPy.

$$VelCPy_n = \frac{CPy_{(n+1)} - CPy_{(n-1)}}{2*t}$$

Where $CPy_{(n+1)}$ = the CPy data point immediately before the nth data point at which velocity is being calculated

$CPy_{(n-1)}$ = the CPy data point immediately after the nth data point at which velocity is being calculated

t = sample rate (0.002 s for this study)

Using CPy error of +/- 0.5 mm (from force plate testing):

Error $= \dfrac{\delta CPy_{(n+1)} - \delta CPy_{(n-1)}}{2*t}$

Maximum error $= \dfrac{0.5+0.5}{2*0.002}$ = 250 mm.s^{-1} = 0.25 m.s^{-1}

Approximate error $= \dfrac{\sqrt{0.5^2+0.5^2}}{2*0.002}$ = 177 mm.s^{-1} = 0.18 m.s^{-1}

Approximate error in 10 trials

$= \dfrac{\sqrt{10*0.18^2}}{10}$ (10 measures each with 0.18m.s^{-1} error)

= 56 mm.s^{-1} = 0.06m.s^{-1}

While the single velocity measures seemed high relative to the measures obtained in this study, the mean value across the 10 trials was low.

Using the theoretical resolution provided by 16-bit ADC (calculations for these values are reported in Ball, 1999)

Theoretical error in CPy using a 16-bit ADC system and using a 700 N golfer is +/- 0.04 mm. Using this value and the equations presented above:

Maximum error	= 20 mm.s^{-1}	= 0.020 m.s^{-1}
Approximate error	= 14 mm.s^{-1}	= 0.014 m.s^{-1}
Approximate error across 10 trials	= 4 mm.s^{-1}	= 0.004 m.s^{-1}

In all cases, the theoretical error estimate of velocity was small.

CPy Velocity at eight swing events

Table E.10 and E.11 report the error in CPy velocity at each swing event using the 200 Hz and 250 Hz cameras. Error was calculated as:

Error due to foot position/force plate data (CPy velocity) + error due to timing (at each event)

Approximate error calculated using quadrature summation while maximum error calculated as the sum of both errors.
Recalling
Error due to force position/force plate data (CPy velocity) = +/- 0.18 $m.s^{-1}$ (maximum = 0.25 $m.s^{-1}$)
Error due to timing from data in tables E.8 and E.9 (note: values were rounded for presentation but not calculation so may seem to be unusual for some calculations).

Table E.10: Error in CPy velocity at each swing event (200 Hz camera).

All values +/-$m.s^{-1}$.

	Approximate		Maximum
	Single measure	Over 10 measures	Single maximum error
TA	0.18	0.05	7.00
MB	0.18	0.04	7.00
LB	0.18	0.04	7.00
TB	0.18	0.06	7.00
ED	0.23	0.15	7.01
MD	0.20	0.10	7.00
BC	0.18	0.04	7.00
MF	0.23	0.16	7.00
AVERAGE	0.20	0.08	7.00

Table E.11: Error in CPy Velocity at each swing event (250 Hz camera).

All values +/- $m.s^{-1}$.

	Approximate		Maximum
	Single measure	Over 10 measures	Single maximum error
TA	0.18	0.05	7.00
MB	0.18	0.04	7.00
LB	0.18	0.04	7.00
TB	0.18	0.05	7.00
ED	0.21	0.12	7.00
MD	0.19	0.09	7.00
BC	0.18	0.06	7.00
MF	0.21	0.13	7.00
AVERAGE	0.19	0.07	7.00

Based on the approximate error over the ten trials the measure was considered reasonable (mean of approximately 0.08 $m.s^{-1}$for each event) and its use in study 2 was considered appropriate.

However, as single trial measures were used in study 3, error may have affected these parameters. While a value of +/- 0.2 $m.s^{-1}$ was indicated as the error for velocity at swing events (and was used in this study) it was probably an overestimation of the error. Firstly, it did not take into account smoothing (twice from the point at which error was calculated here to the value). As well, theoretical analysis suggested the error would be much smaller than this. The true error probably lay between these values. However, correlation analyses were affected minimally when velocity values were rounded to 0.4 $m.s^{-1}$ units, returning similar r-values than when no rounding was performed. As such, this researcher was confident that the data could be rounded to units of 0.4 $m.s^{-1}$, which was possibly an underestimation of the true resolution without generating type 1 errors.

Other CPy Parameters

Timing data

All timing data was accurate to +/- half the sample rate of the force plate in addition to +/- half the sample rate due to possible trigger error (discussed previously).

Maximum	$= 0.001 + 0.001$	= 0.002 s
Approximate	$= \sqrt{0.001^2 + 0.001^2}$	= 0.001 s
Approximate error across 10 trials	$= \dfrac{\sqrt{10 * 0.001^2}}{10}$	< 0.001 s

MaxCPy% MinCPy%

Maximum and minimum CPy% error was the same as for CPy% calculated above as no swing event data was included in these parameters:

Maximum	= 1.60 %
Approximate	= 0.56 %
Approximate error across 10 trials	= 0.18 %

VMaxCPy

As for CPy velocity calculated above.

Maximum error	= 20 $mm.s^{-1}$	= 0.020 $m.s^{-1}$
Approximate error	= 14 $mm.s^{-1}$	= 0.014 $m.s^{-1}$
Approximate error across 10 trials	= 4 $mm.s^{-1}$	= 0.004 $m.s^{-1}$

CPyR

CPyR = CPy Maximum - CPy minimum

δCPyR = δCPy Maximum + δCPy minimum

Maximum	$= 0.5 + 0.5$	= 1.0 mm
Approximate	$= \sqrt{0.5^2 * 0.5^2}$	= 0.7 mm
Approximate error across 10 trials	$= \dfrac{\sqrt{10 * 0.7^2}}{10}$	= 0.2 mm

CPyR%

CPyR% = CPy% Maximum – CPy% minimum

δCPyR% = δCPy% Maximum + δCPy% minimum

Maximum	= 0.56%+ 0.56%	= 1.12%
Approximate	$= \sqrt{0.56^2 * 0.56^2}$	= 0.79%
Approximate error across 10 trials	$= \dfrac{\sqrt{10 * 0.79^2}}{10}$	= 0.25%

Appendix F

Agglomerative schedules and dendrograms for Replication subsets

Replication Subset 1

Table F.1: Selected sections of the agglomerative schedule for hierarchical cluster analysis of CPy% at eight swing events for subset 1 (*N* = 41 golfers).

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	41	153	-	21	21	899	73
2	40	229	75	22	20	932	34
3	39	254	25	23	19	1002	70
4	38	282	28	24	18	1047	46
5	37	292	11	25	17	1052	5
6	36	296	4	26	16	1062	10
7	35	351	55	27	15	1143	80
8	34	414	63	28	14	1307	164
9	33	468	54	29	13	1336	30
10	32	483	14	30	12	1412	76
11	31	491	8	31	11	1507	95
12	30	498	7	32	10	1541	35
13	29	531	32	33	9	1683	142
14	28	583	52	34	8	1778	95
15	27	645	62	35	7	2116	**337**
16	26	710	66	36	6	2126	10
17	25	791	81	37	5	2718	**592**
18	24	801	10	38	4	3528	**809**
19	23	804	3	39	3	4175	**648**
20	22	825	21	40	2	4903	**808**

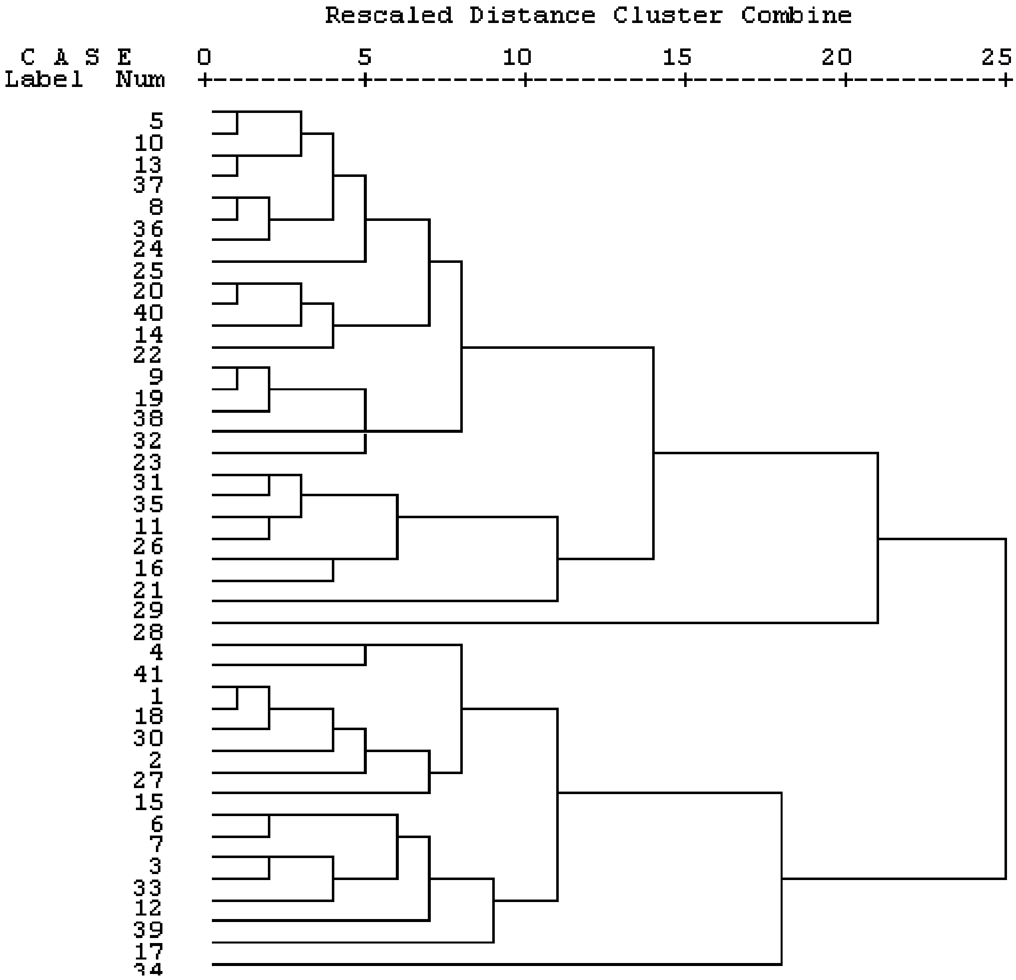


Figure F.1: Dendrogram for hierarchical cluster analysis of CPy% at eight swing events for subset 1 (*N* = 41 golfers).

Replication Subset 2

Table F.2: Selected sections of the agglomerative schedule for hierarchical cluster analysis of CPy% at eight swing events for subset 1 (*N* = 41 golfers).

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	41	142	-	21	21	800	72
2	40	150	8	22	20	824	25
3	39	172	22	23	19	878	54
4	38	252	80	24	18	939	60
5	37	276	24	25	17	984	46
6	36	317	41	26	16	1021	37
7	35	398	80	27	15	1041	20
8	34	408	11	28	14	1068	27
9	33	435	27	29	13	1299	231
10	32	447	12	30	12	1317	18
11	31	457	10	31	11	1400	83
12	30	478	21	32	10	1433	33
13	29	487	9	33	9	1562	129
14	28	515	29	34	8	1707	145
15	27	542	26	35	7	1842	**135**
16	26	603	61	36	6	2092	250
17	25	637	34	37	5	2209	**117**
18	24	681	44	38	4	2657	**448**
19	23	725	44	39	3	3793	**1537**
20	22	727	3	40	2	5326	**1532**

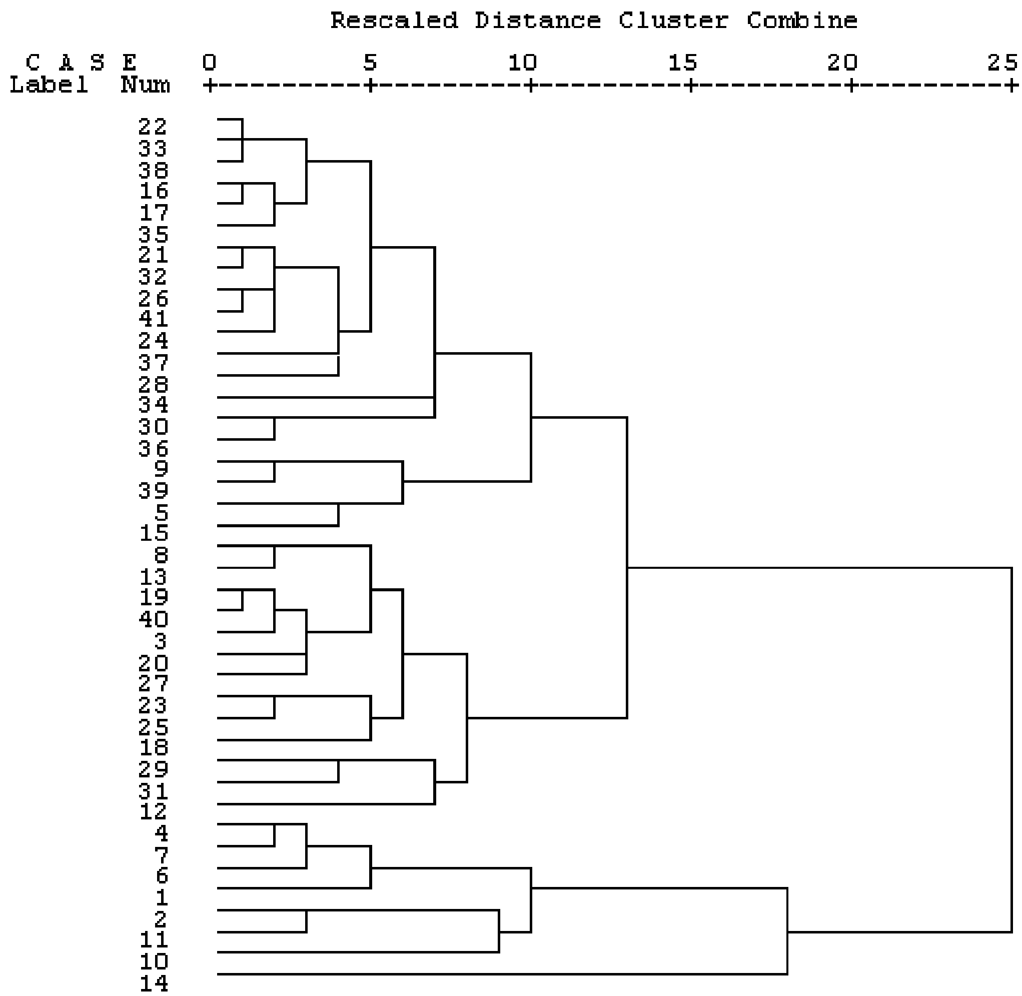


Figure F.2: Dendrogram for hierarchical cluster analysis of CPy% at eight swing events for subset 2 (*N* = 41 golfers).

Replication Subset 3

Table F.3: Selected sections of the agglomerative schedule for hierarchical cluster analysis of CPy% at eight swing events for subset 1 (*N* = 41 golfers).

Stage	Cluster Solution	Coefficients	Jump in Coefficient	Stage	Cluster Solution	Coefficients	Jump in Coefficient
1	41	146	-	21	21	800	24
2	40	224	78	22	20	827	28
3	39	271	47	23	19	845	18
4	38	291	21	24	18	851	6
5	37	300	9	25	17	894	43
6	36	301	1	26	16	969	74
7	35	341	40	27	15	1065	97
8	34	423	81	28	14	1073	8
9	33	457	34	29	13	1191	118
10	32	460	3	30	12	1317	126
11	31	487	27	31	11	1558	241
12	30	521	35	32	10	1563	5
13	29	531	10	33	9	1607	44
14	28	542	11	34	8	1820	213
15	27	607	65	35	7	1821	1
16	26	681	73	36	6	2356	**535**
17	25	703	22	37	5	2461	**105**
18	24	731	29	38	4	2593	**133**
19	23	774	42	39	3	3458	**865**
20	22	776	3	40	2	5542	**2084**

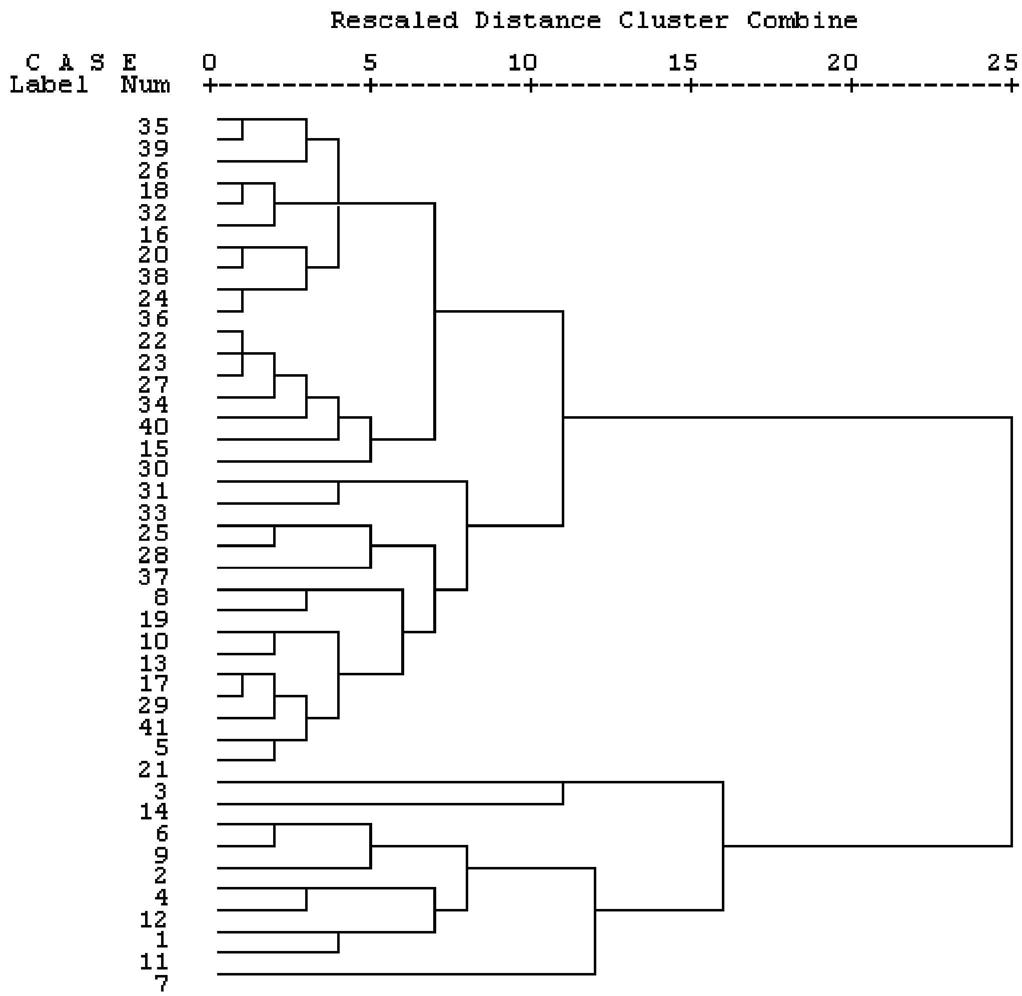


Figure F.3: Dendrogram for hierarchical cluster analysis of CPy% at eight swing events for subset 3 (*N* = 41 golfers).

Appendix G

Poincare plot calculations

To calculate quantitative parameters, the line of identity (x = y, termed P1) was established for the CPyR Poincare plot. A second line was also established – the line perpendicular to the line of identity that passed through the mean CPyR value. This line was referred to as P2.

The perpendicular distance from each (N, N+1) datapoint to the line of identity (P1) was calculated using the following:

Perpendicular distance from P1 for datapoint n

$$= (X_n - Y_n) * \text{Cos } 45^0 \qquad \text{Equation G.1}$$

- Short term variability was calculated as the standard deviation of the perpendicular distances from P1.
- The width histogram was calculated from the perpendicular distances from P1.
- The length of the short axis of the ellipse was calculated as four times the standard deviation of perpendicular distances from P1 (i.e. two standard deviations from the mean in both positive and negative directions - 95% of all values will lie within ± two standard deviations from the mean).

Perpendicular distance from P2 for datapoint n

$$= (Y_n + X_n - 2 * \text{mean CPyR}) * \text{Cos } 45^0 \qquad \text{Equation G.2}$$

- Long term variability was calculated as the standard deviation of the perpendicular distances from P2.
- The Length histogram was calculated from the perpendicular distances from P2.
- The length of the long axis of the ellipse was calculated as four times the standard deviation of perpendicular distances from P2 (i.e. two standard deviations from the mean in both positive and negative directions - 95% of all values will lie within ± two standard deviations from the mean).

Proofs for equations 7.1 and 7.2 are provided below.

The 95% area ellipse (adapted from AMTI, 1982)

$$= \pi * \frac{Length(ShortAxis)}{2} * \frac{Length(LongAxis)}{2} \quad \text{Equation G.3}$$

Where

- *Length(ShortAxis)* = four times the standard deviation of all perpendicular distances from P1 (two standard deviations either side of line P1 encompass 95% of the datapoints hence the 95% area ellipse).
- *Length(LongAxis)* = four times the standard deviation of all perpendicular distances from P2 (two standard deviations either side of P2 encompass 95% of the datapoints hence the 95% area ellipse).

Proofs for equations G.1 and G.2

Perpendicular distance to the line of identity

Figure G.1 is the reference figure for the calculation of perpendicular distance from point A to the line of identity.

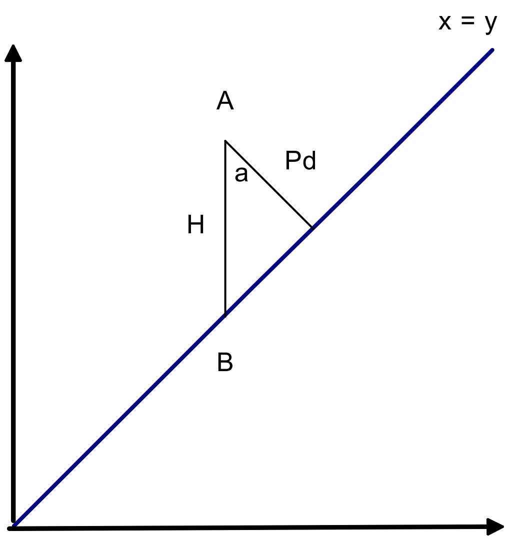


Figure G.1: Example calculation of perpendicular distance from point A to the line of identity

From Figure G.1

Length H = Ay – By

Where
Ay = current datapoint (value known)

By = vertical projection of A to line X = Y

Therefore By = Bx
And Bx = Ax (vertical projection from A means x-coordinate does not change)
So
By = Ax

Thus

H = Ay - Ax
Using Trigonometry and referring to figure G.1

Perpendicular distance Pd from point A to line x = y

$$Cos.a = \frac{adjacent.side}{Hypotenuse}$$

Rearranging

Adjacent side = Hypotenuse * Cos a

Or

Pd = H * Cos a

Where a = 45^o (x = y at 45^o to vertical and line H is vertical)

So

Pd = (Ay – Ax) * Cos 45^o

Or in general terms, for datapoint n with coordinates (X_n, Y_n), the perpendicular distance to the line of identity is given by

(X_n – Y_n) * Cos 45^0

Perpendicular distance to a line perpendicular to the line of identity and passing through the mean parameter value.

Figure G.2 is the reference figure for the calculation of perpendicular distance from point A to line perpendicular to the line of identity and passing through the mean parameter value

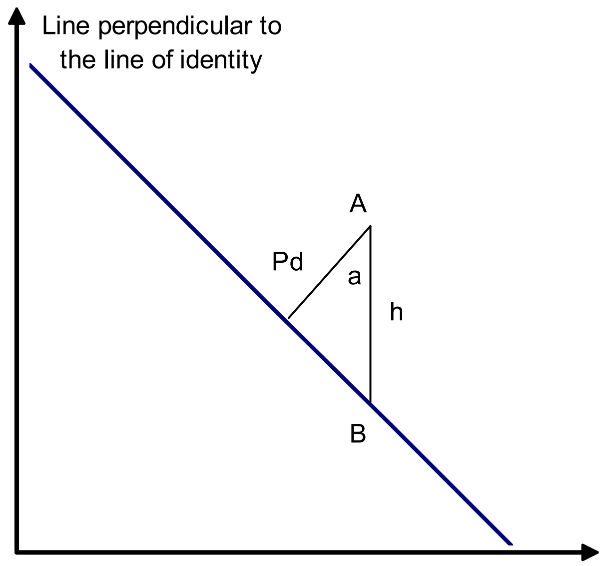


Figure G.2: Example calculation of perpendicular distance from point A to a line perpendicular to the line of identity.

From Figure G.2

$y = mx + b$

The slope of the line is known but the y-intercept is not
$m = -1$

rearranging

$$\begin{aligned} b &= y - mx \\ &= y + x \end{aligned}$$

Setting this line to pass through the mean value of the parameter being examined (meanParameter)

$$\begin{aligned} b &= \text{meanParameter} + \text{meanParameter} \\ &= 2.\text{meanParameter} \end{aligned}$$

So

$$y = -x + (2.\text{meanParameter})$$

Referring to figure G.2

h $\quad = Ay – By$

Where
Ay $\quad$ = current datapoint (value known)
By $\quad$ = vertical projection of A to line y = -x + 2.meanParameter
$\quad$ = -Ax + 2.meanParameter

Thus

h $\quad$ = Ay – (-Ax + 2.meanParameter)
$\quad$ = Ay + Ax – 2.meanParameter

Using Trigonometry and referring to figure *
Perpendicular distance Pd from point A to line y = -x + 2.meanParameter

$$Cos.a = \frac{adjacent.side}{Hypotenuse}$$

Rearranging
Adjacent side = Hypotenuse * Cos a
Or
Pd = h * Cos a
Where a = 45^o
So

Pd = (Ay + Ax – 2.meanParameter) * Cos 45^o

Or in general terms, for datapoint n with coordinates (X_n, Y_n), the perpendicular distance to a line perpendicular to the line of identity and passing through the parameter mean is given by

$$\mathbf{(Y_n + X_n - 2 * mean\ CPyR) * Cos\ 45^0}$$