Inter-Rater Reliability and Criterion Validity of Scatter Diagrams as an Input Method for Marksmanship Analysis: Computerised Notational Analysis for Archery

Andrew J. Callaway and Shelley A. Broomfield

Centre for Event & Sport Research, Bournemouth University, Talbot Campus, Bournemouth. BH12 5BB, UK.

Abstract

Basic notational systems exist in many sports but may not always be formally named as such. This paper outlines the validity and reliability of using a computerised scatter diagram as an input method for the analysis of marksmanship in target sports with Archery used as an example sport. Software was created to represent a standard FITA 122cm target face. An Archer shot 72 arrows, these were measured to create coordinates of each arrow. Twelve images (6 arrows per image) were shown to 10 operators. Operators input each arrow position into the software to produce a coordinate for each arrow. The coordinates where additionally processed using marksmanship statistics, Accuracy Constant Error (A_{CE}), Accuracy Variable Error (A_{VE}) and Precision (Mean Radius- MR). ICC was used to test validity and PPMC to test inter-rater reliability of the raw data and error analysis on the processed data. Raw results showed strong positive correlations for validity against the measured coordinates (>.977) and high levels of inter-rater reliability(>.969). Processed results show less than 2% error to that of measured results. Systems such as this allow coach/athlete/researcher to track varying equipment setups, changes to biomechanics, physiology and psychology, allowing for continued development of athlete, sport and equipment.

Keywords: Notational Analysis, Reliability, Validity, Shooting, Marksmanship

1. Introduction

Notational analysis has been used in sports to characterise factors which determine success of teams and individuals, be these tactical, technical or technique aspects or to record performances over time for comparison (Hughes and Franks, 2004). Methods of recording this information varies between manually tallying occurrences to noting where an event occurs on a schematic layout of the playing surface (or schematic of an athlete); whichever is appropriate to task (Hughes and Franks, 2004); through to computerised systems such as Prozone where players can be tracked (Di Salvo *et al.*, 2006).

Notational systems can exist in many sports but may not always be formally named as such, where Archery is one such example. Competitions, from local events to the

Olympics all rely on the accumulation of points, and this therefore is the most popular criterion to analyse the performance of an archer in studies (Hay and Reid, 1988; Leroyer *et al.*, 1993; Lin and Hwang, 2005). Archers shoot rounds (a combination of distances shot, target face size and scoring zones) which can vary depending on the age group, gender and bow discipline (World Archery, 2010a; 2010b).

Typically research in archery uses FITA rounds (Keast and Elliot, 1990; Ertan *et al.*, 2005), using a 10 zone scoring system where scores range from 10 to 1 with a miss scored as "M". The centre, 10 ring, is separated into two sections, which are recorded as 10 for the outer section and X for the inner, however for addition purposes, both score 10 points. The rationale is to determine the ranking position for archers on the same score; same score with a higher number of X's wins the higher place. These scores will then be used to rank the archers for medals or in events such as the Olympics, for the elimination rounds (Head to Head matches).

The recorded scores are typically written on a sheet of paper, with some high level competitions using a computerised score entry pad, where each end (set) of arrows shot are recorded in order of highest to lowest value and added for a cumulative total, effectively a form of frequency table. This level of data input has limitations in the degree of analysis that can be performed on that data as the arrows are only scored from highest to lowest, not even in the order shot. Scatter diagrams are a development from this and appear in previous research to locate setting zones in volleyball (Koch & Tilp 2009), actions in Netball (Bruce *et al.* 2009), representing a goal in football (Bar-Eli and Azar, 2009; Van der Kamp, 2011) and additional examples demonstrated by Hughes and Franks (2004). A scatter diagram approach in archery allows arrow positions to be plotted onto a target face, allowing the archer to see personal best records for each tournament showing average groupings, individual arrow performance and grouping by distance. Systems for personal use tend to appear as scatter diagram based software for PDAs (Personal Digital Assistants), such as BowBuddy (Bowbuddy, 3.4) with alternatives starting to appear for the smart phone market on varying operating systems. Software such as this can be used by the archer/coach to record a more detailed performance over score.

Feedback and information from these systems in Archery can be limited to a visual representation of arrows in a target showing a group which allows athletes to make immediate changes to equipment.

Using a computerised system utilising arrow locations (coordinates) based on real measurements also allows for a greater level of analysis such as marksmanship statistics for Accuracy Constant Error, shown in equation (1), Accuracy Variable Error in equation (2) and Precision shown in equation (3) (Johnson, 2001). This can then allow coaches, researchers and archers, investigating aspects such as the effect of biomechanical changes on outcome, equipment changes, as well as physiological and psychological factors, require a method to gauge in changes in *real* terms, such as centimetres and more developed than just score.

$$A_{CE} = \sqrt{(\bar{x})^2 + (\bar{y})^2} \qquad\qquad\qquad (1)$$

$$A_{VE} = \frac{\sum \sqrt{(x_i)^2 + (y_i)^2}}{n} \tag{2}$$

$$MR = \frac{\sum \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}}{n} \tag{3}$$

These statistical measures would be questionable without a valid input by the user (Wilson and Barnes, 1998). The aim of this study was to determine the criterion validity and inter-operator reliability of inputting "arrows" into a computerized system to allow coaches, researchers and maybe some archers, to better understand the outcome of the arrows on a target. As this system has the potential to be used for more than just recording scores, the marksmanship statistics were also tested as the equations each sum distances (coordinates), and therefore sum any input error.

2. Method

2.1. Study design
With institutional ethical approval, an archer shot 72 arrows (FITA 70m) at 70m on a standard 122cm target face. Each end of 6 arrows was photographed and on pulling each arrow from the target the coordinate of each arrow was measured (x and y), to the centre of each hole left by the arrow, to the nearest millimetre. This produced master coordinates of each arrow in the order shot by the archer.

The photographs were taken with the centre of the lens of the camera standing 130cm, 90 degrees to the target and 80cm from the target. The camera used was a Nikon Coolpix S3000X 12MP Camera. No zoom was used to ensure repeatability. The photographs were not manipulated apart from to add labels to show the order in which they were shot, 1-6 (Figure 1 a).

A Matlab script was developed, which superimposed a blank target face over a set of axis (MatLab, 2007) (Figure 1b). The centre of the target was set to (0,0) and the sizes of each colour band were in accordance to the limits set by World Archery (2010b) for a 122cm target. The software can be altered to replicate other target sizes and with a zoom control available, changing this target size in the software would not affect the visual representation to the user, only the internal calculations would be affected. An arrow position was recorded in the software by a click, which generated a coordinate (x,y) recorded in centimetres to an accuracy of 4dp. The software then took the 6 arrows entered and calculated the Accuracy Constant Error (A_{CE}) (1), Accuracy Variable Error (A_{VE}) (2) and Precision Mean Radius (MR) (3). The same statistics were also used for the full 72 arrows once all were entered. The master coordinates were also processed in this manner to allow a comparison to operators' raw input error as well as the error after these calculations.
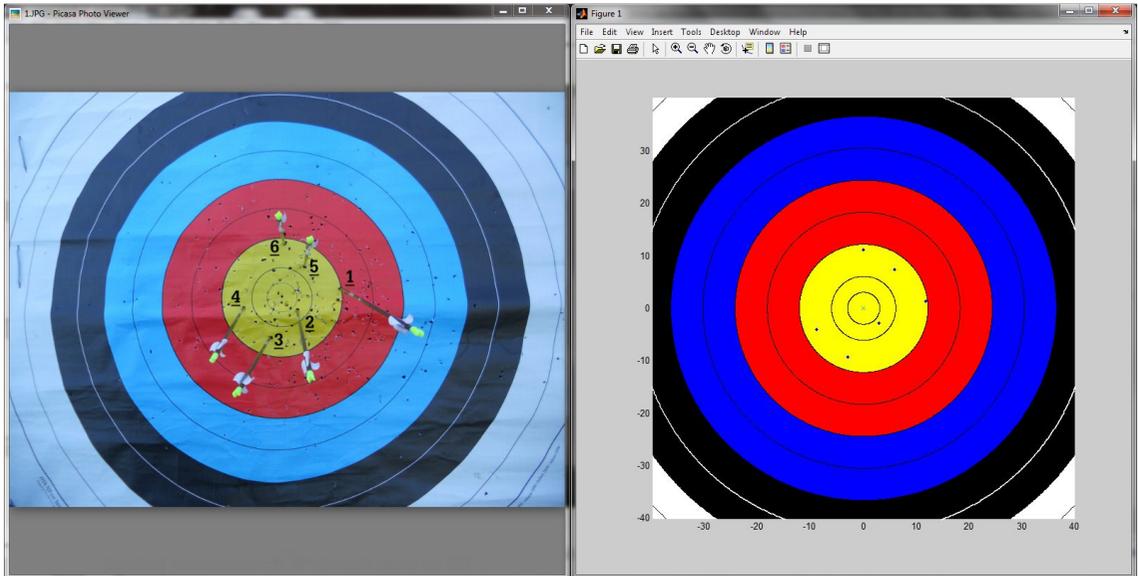
Figure 1. Computer system set up. a) Photograph of the target face with labelled arrows. b) Computerised target face for input from user

2.2. Participants
Ten participants provided informed consent to take part in this experiment. Knowledge of Archery was not deemed necessary for this research. Seven male and three female participants took part (29.5yrs ±7.6yrs).

2.3. Procedure
The objective was for the operator to input the arrow locations from the photograph into the software to the best of their ability. The software was fully demonstrated to each user separately. Zoom controls on both the photograph and the software were made clear to allow for personal preference. With a blank target shown on the software, they were instructed to click, to the best of their ability, the location of each arrow in numerical order on the photograph. Once all 6 arrows from the photograph were input, the target on the software was cleared and the next photograph was shown and process was repeated for 12 photographs (72 arrows in total). Each operator had 4 minutes to complete each set of 6 arrows, to represent the same timings as in a qualification round at an event. Each operator completed each end well within these time limits.

2.4. Statistical analysis
2.4.1. Scores
Score is the primary function of the notational system, errors in scores are analysed as a % difference from the actual score of from the archer (649), using equation (4).

$$Score \ \% \ error = \ 100 \left(\frac{Actual \ Score - Operator \ Score}{Actual \ Score} \right) \tag{4}$$

2.4.2. Validity and Reliability of Raw Data Entry

Typically within validity and reliability studies the actual results are not known, so are often compared to a previous "gold standard". This study's gold standard is the measured coordinates themselves allowing measurement of criterion validity; as a subdivision of concurrent validity (Wilkinson *et al.*, 2009; Bannigan and Watson, 2009; O'Donoghue, 2010). This can be measured using Pearson Product Moment Correlations (PPMC) of each operators input against the master coordinates. This was repeated for the abscissa (X) and ordinate (Y) inputs separately.

Inter-rater reliability was assessed using Intra-class Correlation Coefficient (ICC) where each operator was compared, and repeated for the abscissa (X) and ordinate (Y) inputs separately (Shrout and Fleiss, 1979; Weir, 2005).

The difference between each operators' input and the master input were also analysed using Mean Error (±SD), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) for each (Hughes *et al.,* 2002; O'Donoghue, 2010; 2012). Systematic and random bias are also reported for each axis.

The marksmanship statistics, A_{CE}, A_{VE}, MR, all rely upon the straight line distances from the centre point. Therefore, error at this stage of the process is of great importance, as significant error in the input of the data would filter through to the output. Hughes, Cooper and Nevill (2002) and O'Donoghue (2010) state it is important to explain the method of percentage error used. There are 2 forms in which this straight line error can be realised (Figure 2). Firstly by the error on either axis (Figure 2, a and b) or the straight line distance from the centre (Figure 2 c). Angular deviation was not considered here as the marksmanship statistics are more concerned with the straight line distance to represent group centres. The mean %error for the abscissa (x) and ordinate (y) can be calculated using the modulus of the master coordinate (X) less the operators input (*x*) divided by the targets radius shown in equation (5). This can be summed for each coordinate in the system, and multiplied by 100 to return a mean %error for the abscissa and ordinate.

$$Mean \: \%error = \: 100 \left(\sum_{i=0}^{N} \frac{|X-x|}{TargetRadius} \right) \tag{5}$$

Developing this, the error of the straight line distance from the origin, of these inputs can be tested (Figure 2 c). This will be named, Straight Line Mean Percentage error (SLM %error). Pythagoras theorem was used to calculate the straight line distance of the Master Coordinates (*X, Y*) and the operators input (*x, y*). Taking the modulus of the master coordinates less the operators and dividing by the radius of the target will result in a SLM %error. This process can also be summed to give a SLM %error for all the inputs shown in equation (6). ICC and PPMC were calculated using SPSS version 19 (IBM, 2010), all other calculations were calculated using Excel (Microsoft, 2007).
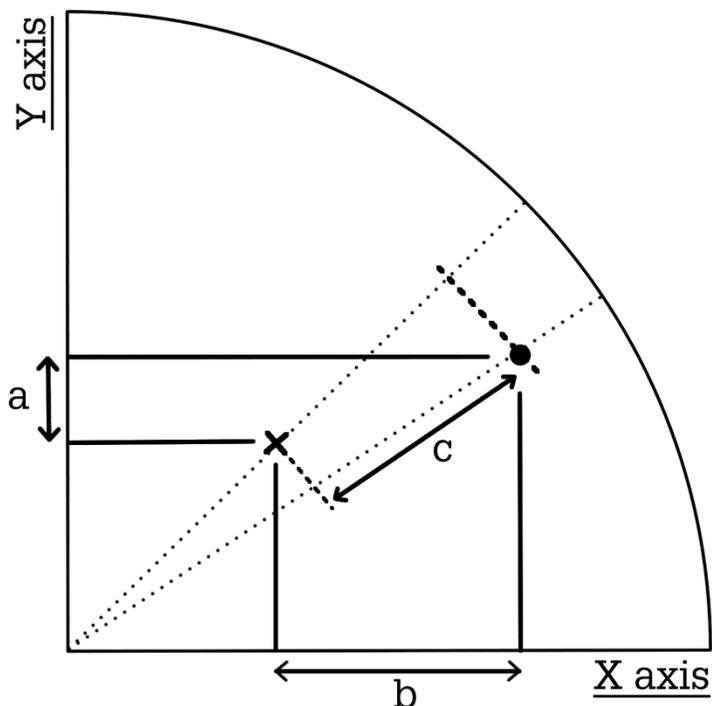
Figure 2. A quadrant of the target demonstrating the ways in which error can be seen on arrow placement. a and b show error between X (actual) and ● (operator placement) on each axis. c shows error as a measure of straight line distance from the centre of the target (Pythagorean distance). The outer circle represents the circumference of the target face with the axis being the diameter of the target face.

$$SLM\ \%error\ =\ 100\left(\sum_{i=0}^{N}\frac{\left|\left(\sqrt{X_i^2+Y_i^2}\right)-\left(\sqrt{x_i^2+y_i^2}\right)\right|}{TargetRadius}\right) \tag{6}$$

2.4.3. Validity and Reliability of Processed Data

The raw input, as discussed, can be used with A_{CE} (1), A_{VE} (2), MR (3). The equations use sums of locations, and therefore can sum error, this needs to be addressed in order to ascertain its validity and reliability. The measured and operator input coordinates were processed using A_{CE}, A_{VE}, MR. Validity was then assessed using PPMC against each operator. Inter-rater reliability was assessed using ICC.

The difference between each operator's results and the master results will be shown using Mean Error (±SD), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and %error for each.

Percentage error for A_{CE} and A_{VE} was calculated using the same process as (6), replacing the distance term (X) with the master A_{CE} less the operators A_{CE} (x) (or A_{VE} as appropriate). The mean %error is reported for A_{CE} and A_{VE}.

Percentage error of the Mean Radius (MR) was calculated as the difference in the area of the circle produced by the master coordinates (MR) and the radius produced from each operator (mr) and divided by the target area, and multiply by 100, in equation (7).

$$Radial\ \%error = \ 100 \left(\sum_{i=0}^{N} \frac{\left|(\pi(MR)^2)-(\pi(mr)^2)\right|}{\pi TargetRadius^2}\right) \tag{7}$$

Finally, a visual representation of the output from the system is also demonstrated on a target face of the first end (first 6 arrows) and all 72 arrows. This was calculated using the coordinates, from any operator, with the minimum and maximum error. The A_{VE} calculation determined the centre of the group, and the MR determined the size of the group.

3. Results

3.1. Scores
All operators scored 100% accuracy (0% score error) in the score value, except for operators 3 and 8. Operator 3 made 3 errors in the scoring, (+1, -1, -1) and Operator 8 made a single mistake (+1 point). Whilst the primary goal of shooting sports is score, from a total score of 649 these two errors report as 0.3% and 0.15% error respectively. It is felt that experienced operators would further limit this error.

3.2 Raw Data Results

Table 1. Validity of each user correlated against the measured results using PPCM (p = 0.00)

	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10
Y Axis	0.995	0.996	0.993	0.995	0.993	0.993	0.986	0.988	0.986	0.977
X Axis	0.994	0.995	0.987	0.995	0.996	0.994	0.992	0.985	0.991	0.984

Table 2: Inter-Rater Reliability using ICC (p = 0.00)

	U1		U2		U3		U4		U5		U6		U7		U8		U9	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
U2	0.995	0.996																
U3	0.989	0.990	0.986	0.993														
U4	0.996	0.994	0.996	0.996	0.989	0.990												
U5	0.995	0.994	0.996	0.995	0.988	0.988	0.996	0.994										
U6	0.995	0.992	0.996	0.995	0.987	0.984	0.997	0.995	0.995	0.993								
U7	0.995	0.988	0.994	0.987	0.982	0.978	0.996	0.990	0.993	0.989	0.996	0.990						
U8	0.986	0.990	0.986	0.991	0.977	0.982	0.990	0.993	0.986	0.989	0.990	0.989	0.989	0.988				
U9	0.993	0.985	0.993	0.988	0.978	0.981	0.993	0.989	0.994	0.988	0.992	0.988	0.992	0.980	0.985	0.982		
U10	0.986	0.980	0.986	0.981	0.974	0.969	0.987	0.986	0.988	0.986	0.989	0.986	0.989	0.985	0.982	0.979	0.985	0.976

Table 3. A table showing the Mean, MAE, RMSE, Mean %error and SLM %error (6) of the raw data entry from each operator over all 72 arrows.

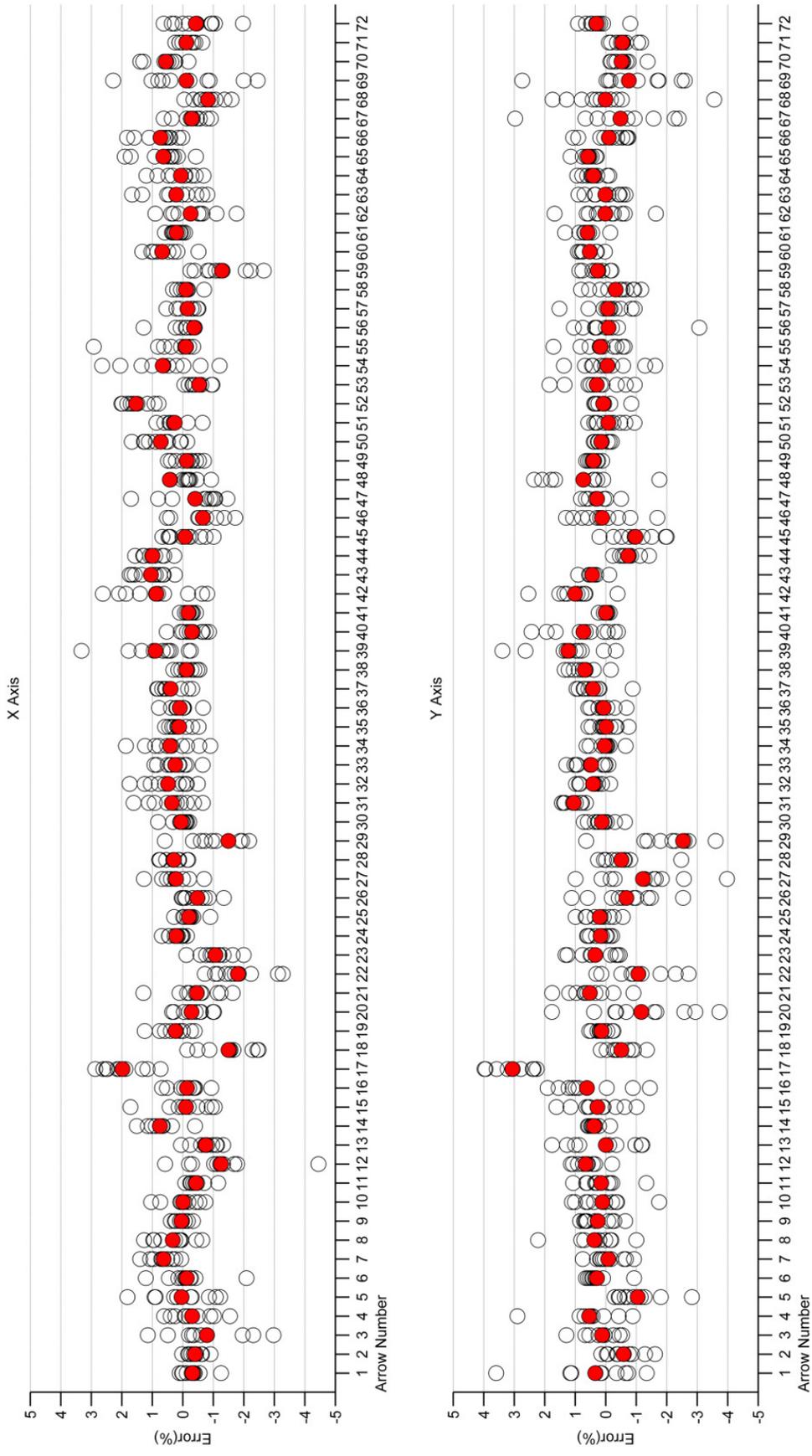| Operator	X Axis				Y Axis				
	Mean (± SD) (cm)	MAE (cm)	RMSE (cm)	Mean % Error (%)	Mean (± SD) (cm)	MAE (cm)	RMSE (cm)	Mean % Error (%)	SLM % Error (%)
1	0.89 (±7.83)	6.00	6.00	0.98	1.72 (±7.16)	5.60	5.60	0.92	0.69
2	-1.15 (±5.01)	5.34	5.34	0.87	0.79 (±4.15)	5.16	5.16	0.85	0.69
3	0.47 (±9.51)	7.88	7.88	1.29	0.95 (±5.34)	6.53	6.53	1.07	0.88
4	-0.75 (±4.81)	5.58	5.58	0.91	1.68 (±4.89)	5.45	5.45	0.89	0.54
5	-0.93 (±4.27)	5.29	5.29	0.87	1.46 (±6.72)	6.20	6.20	1.02	0.71
6	-1.82 (±5.40)	5.87	5.87	0.96	-0.55 (±5.45)	6.61	6.61	1.08	0.76
7	-0.43 (±6.85)	6.43	6.43	1.05	2.08 (±8.63)	7.89	7.89	1.29	0.60
8	1.54 (±9.83)	8.43	8.43	1.38	1.21 (±6.98)	7.54	7.54	1.24	0.91
9	3.96 (±6.58)	8.36	8.36	1.37	-2.44 (±9.01)	9.56	9.56	1.57	0.83
10	-2.99 (±10.08)	10.01	10.01	1.64	1.34 (±12.65)	11.14	11.14	1.83	0.82
Mean	-0.12 (±7.02)	6.92	6.92	1.13	0.82 (±7.12)	7.17	7.17	1.17	0.74

Figure 3. Percentage error for each arrow shown for X and Y axis separately. Red dots (●) represent the mean for all 10 operators. White dots (○) represent the each operators input.

300

The validity of the system is shown in Table 1. Here we can see that the operators input all produce high significant correlations for each axis when compared to the measured results. Operators 7, 8, 9 and 10 produce weaker results, although still significantly high, under the 0.99 level. This could hint towards a weakness with their ability to judge vertical distances, although the correlation is still very high.

Reliability of the system was tested using Inter-Rater reliability using ICC shown in Table 2. This again shows very high levels of internal correlations. Compared to the results in Table 1, we see some additional drops towards the 0.97 mark, however these are still significant and not detrimental to the nature of the system. This demonstrates that the system is reliable between users.

Table 3 shows the mean results for each user. As shown in Table 1, Operators 7, 8, 9, 10 showed lower correlations when compared to the measured results. This is highlighted in Table 3 where operators 8, 9 and 10 show the highest error with Operator 3. Operator 3 shows a large MAE in the X and Y axis, this is not replicated in Table 1, but could be hinted at in Table 2 showing the lower ICC than some of the other operators. However, their standard deviation (SD) about the mean is amongst the highest. The small mean and large SD shows that their positive and negative error could be cancelling itself out, in terms of the calculations used. Operator 7; shown to have lower correlations in Table 1; does not show high errors in the X axis but does in the Y axis. Overall we see that operators tend to under estimate in the horizontal plane (X) placing arrows closer to the centre, yet over estimate in the vertical plane (Y). In spite of this, the results (Table 3) also show that all Operators, over all 72 arrows, have a mean error of less than 1% in the straight line distance from the centre point of the target. So whilst there is more error in the Y axis on average, when combined, the overall resultant error is actually minimised. The RMSE is equal to MAE demonstrating no significant variation in the error. The systematic error for the system is shown to be -0.12, 0.82 for X and Y respectively with the random error being ±7.02 and ±7.12.

Figure 3 shows the error for each of the 72 arrows. This shows that Arrow 17 and 29 show the largest errors. Arrow 17 shows large errors in both X and Y axis. Arrow 29 shows a large error in the Y axis, but comparable in the X axis.

3.3 Processed Results

Table 4. shows a summary of the error between each operator and the master calculations for A_{CE}, A_{VE} and MR.

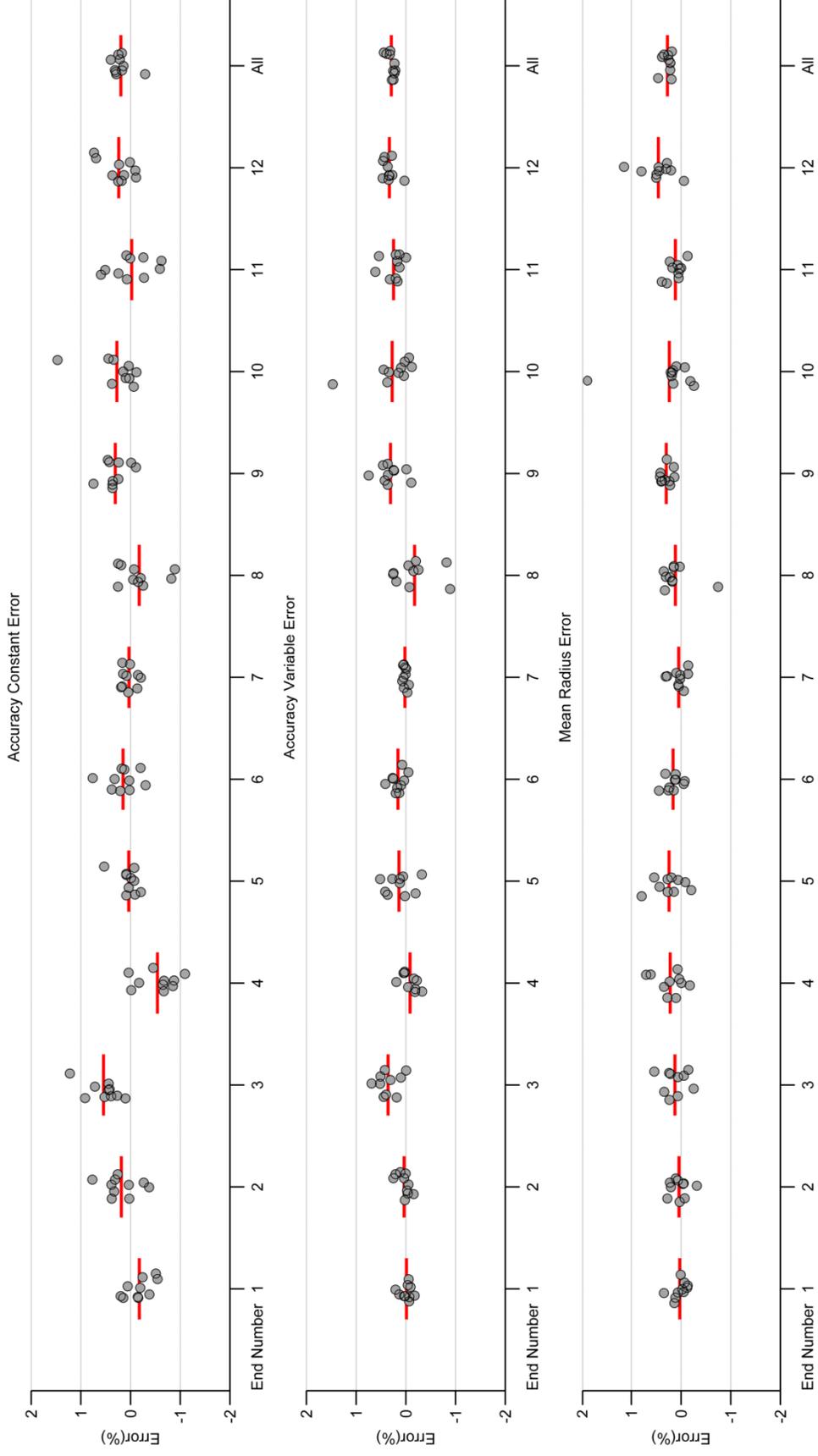| Operator	ACE				AVE				MR			
	Mean (± SD) (cm)	MAE (cm)	RMSE (cm)	SLM %Error	Mean (± SD) (cm)	MAE (cm)	RMSE (cm)	SLM %Error	Mean (± SD) (cm)	MAE (cm)	RMSE (cm)	Radial %Error
1	0.78 (± 2.88)	2.82	3.95	0.46	2.28 (± 4.62)	2.75	5.22	0.45	2.37 (± 3.86)	3.02	4.78	1.28
2	1.24 (± 3.49)	3.08	4.55	0.50	2.13 (± 4.36)	2.80	5.03	0.46	2.09 (± 3.5)	2.40	4.17	1.03
3	2.40 (± 4.92)	2.94	5.55	0.48	4.42 (± 7.28)	4.49	8.29	0.74	4.32 (± 6.46)	4.41	7.60	1.98
4	1.44 (± 3.01)	3.23	4.33	0.53	2.48 (± 4.85)	2.87	5.47	0.47	2.44 (± 3.88)	2.78	4.65	1.19
5	2.58 (± 4.78)	4.16	6.19	0.68	3.88 (± 4.87)	3.88	6.07	0.64	3.26 (± 3.32)	3.26	4.55	1.47
6	4.03 (± 2.99)	4.03	4.95	0.66	2.53 (± 5.15)	3.77	6.21	0.62	2.11 (± 4.37)	3.31	5.34	1.47
7	1.67 (± 3.46)	4.40	5.51	0.72	3.13 (± 4.84)	3.56	5.85	0.58	2.48 (± 4.10)	2.76	4.81	1.29
8	2.34 (± 2.68)	3.50	4.35	0.57	2.29 (± 4.67)	4.31	6.21	0.71	2.28 (± 4.38)	4.22	5.95	1.87
9	-0.67 (± 2.78)	4.51	5.23	0.74	2.79 (± 4.90)	3.45	5.82	0.56	3.98 (± 4.68)	4.58	6.41	1.93
10	2.17 (± 4.52)	5.74	7.20	0.36	3.20 (± 5.27)	3.65	6.23	0.53	3.17 (± 4.13)	4.02	5.64	1.77
Mean	1.52	3.84	5.18	0.57	2.92	3.55	6.04	0.58	2.85	3.48	5.39	1.53

Figure 4. Percentage error plot for A_{CE}, A_{VE} and MR over each end (each set of 6 arrows) and all 72 arrows. The red line (-) show the mean result for the Operators. Each grey dot (•) represents the Operators results.

Whilst in the raw data (Table 3) there was no difference between MAE and RMSE, the processed data (Table 4) shows variance in the error. The greater the difference between them, the greater the variance in the individual errors in the sample, this shows the accumulation of the error from the individuals inputs. Compared to the raw results (Table 3), Operators 7, 8, 9, 10 (Table 4) were not the worst recorded here although still amongst the top of the group. Operator 6 shows the largest Mean error, but MAE and RMSE are fairly similar showing that the error from this Operator is fairly consistent when compared to, for example, Operator 10 who has a far larger discrepancy, showing a more varied deviation from the input data.

There is a varied mean result in A_{CE} for the first 4 ends (Figure 4). This does seem to settle over time so could be due to the novice operators getting acquainted with the system. This variation is not replicated in the A_{VE} or MR results suggesting that these calculations would not be as susceptible to potential erratic errors.

As a demonstration of the effect of the error in a users input to the system, Figure 5 shows the calculation of the Accuracy (A_{VE}) and Mean Radius (the outer coloured circle about the A_{VE} marker) using the minimum value (a and e) and the maximum value (b and f) entered by any of the 10 operators, in comparison to the actual measured values (c and g). This was produced to show how little visual difference is created from the range of the inputs recorded. The image is cropped to the 9 ring of the target, measuring 24cm in diameter.
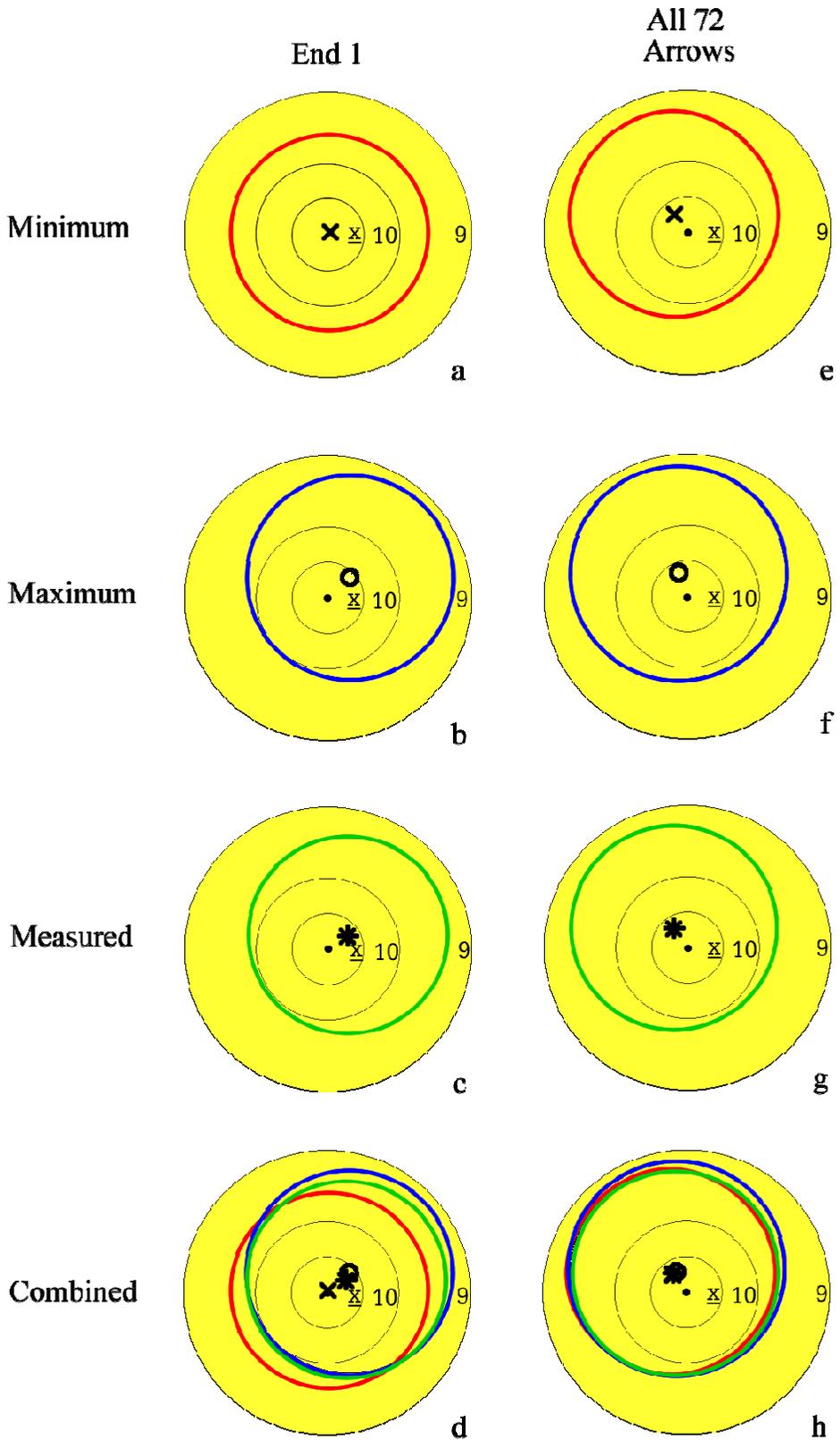
Figure 5. Representation of error using Minimum AVE (x) with Minimum Mean Radius (**red**) and Maximum AVE (o) with Minimum Mean Radius (**blue**) and measured AVE (*) with measured Mean Radius (**green**) from all 10 Operators ($\underline{X}$, 10, 9 indicate the scoring zone, ● indicates the target centre).

4. Discussion

The results demonstrate the system to be reliable and valid between users and when compared to the measured coordinated. We do see an over estimation in a vertical input (Y axis) to the system. This has also been shown to be the case with research by Dixon and Bird (1981) when using a vertical 10cm scale to measure pain, compared to a horizontal scale. This perception error as has also been shown in more complex virtual environments by Vishton *et al.* (1999) and Plumert *et al.* (2005), showing overestimation in the virtual environment compared to real world. Within the current system, this error could be incorporated into as a calibration routine for each user where they input X arrows into the system to generate a X and Y bias to remove/add to their remaining data input, in the same way that users calibrate touch screen computers (Vidales, 2002).

Perception error in general notations systems warrants further investigation to see if procedures or designs of systems can be altered to help reduce this error. This could include shapes, colours, shading and reference points. An example of reference points within existing systems, primarily court/pitch based games, are likely to be the court lines. When entering the data into a notational system, Hughes and Franks (2004) show that by dividing up the space into a smaller area allows for a more detailed analysis, but choosing the correct 'box' to put a position in then becomes harder and a potential source of error. This all links to the perception errors within notational analysis which could be investigated further. Within Archery and other shooting sports, the targets are typically banded colours which limits the use of reference points other than the target centre, so previous arrow positions could be used as a reference.

Error identified within the present system shows arrow 17 and 29 to be problematic. The images show an overlap between arrows 5 and 6 (Figure 6a, Arrow 5), occluding the point of entry resulting in a large error from the operators. The image containing arrow 29 has no occlusion of the arrow (Figure 6b, Arrow 5), however, the entry sequence of the two preceding arrows may have influenced the result.
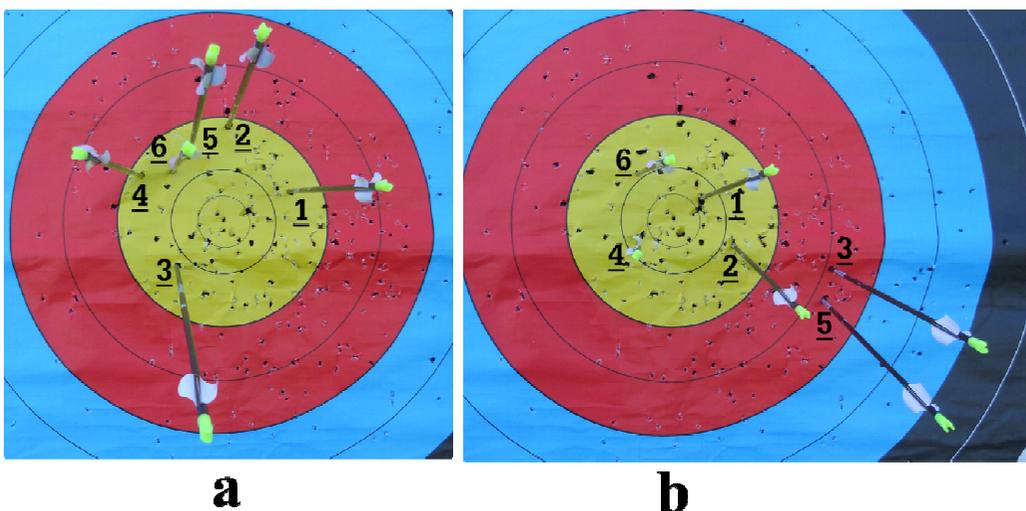

Figure 6. a) Arrow 17 (arrow 5 on image) b) Arrow 29 (arrow 5 on image) (Cropped image)

The two previous arrows were outer red (Figure 6b, Arrow 3) and inner gold (Figure 6b, Arrow 4). If an error was made with the placement of Arrow 3, then the Operator may have used this as a reference point, magnifying the error. Errors such as this are unavoidable using the present system, as is occlusion of any arrows entry point into the target. An archer using this system on a PDA/Tablet at the target, could have an advantage over a coach at the side line. The archer could move around close to the target to get the best view possible, before entering the data into this type of system.

Hughes *et al.* (2002) demonstrate errors in entering the wrong information showing that nearly 3% of their errors were due to inputting the wrong cell number. This shows favourable results where only 2 operators had these errors; arrows input into the wrong scoring band; totalling error of 4.1% (Operator 3) and 1.3% (Operator 8). Hand notation systems will inevitably have a degree of error due to the humans entering the data. Hughes, Cooper and Nevill (2002) analysed the number of shots in live Squash games using two analysts. Their results demonstrated less than 1% error when recording number of shots and was then developed into location ball-wall contact, with the wall divided into 16 cells. With this progression, the researchers predicted a higher level of error from counting shots and demonstrated error of over 30%. Choi *et al.* (2007) showed inter-operator reliability between 4 analysts recording variables in basketball, their results demonstrated errors between analysts from 14.8% to 33.8%. Any issues with previous hand-entry systems are most likely due to the speed of the games analysed. The longer period of 4 minutes to analyse the arrow locations may explain the lower error seen in this study compared to others.

Maslovat and Franks (2007) discuss that the use of video and computer based technologies can allow for comprehensive post-event analysis, where information can be slowed to review. Within archery slowing videos would be necessary for biomechanical analysis, yet for recording arrow scores, this would not be necessary as the scoring process is currently fairly slow.

Arrows are currently scored when all archers have finished shooting and everyone will approach the targets. There is scope for a camera based system, or laser based system to record these arrow positions either as they happen, or after completion of the set (end) of arrows however there are potential cost implications which might only be useable at top level competitions limiting the analysis potential for up and coming athletes.

Danage (2012) produce a scoring system where the archers, when at the targets, enter the scores of the arrows in order of highest to lowest. Whist entering arrow scores electronically like this offers a simple solution for summation and the ranking of individuals in competition or at the club in practice, there are limitations in the degree of analysis that can be performed on that data due to a lack of information gathered when entering the data. Other implications for computerised systems are when multiple archers are shooting at the same time, with typically 3-4 people per target, all with different colour combinations of fletching, arrow wraps and nocks these could induce further occlusion of the arrows. Where a person can quickly see the difference and entry point, a computer could have difficulty with some of the occlusions, especially with top archers shooting very tight groups at a short distance.

Computerised systems that appear for personal use, tend to appear as scatter diagram based software for PDAs (Personal Digital Assistants), such as BowBuddy (Bowbuddy, 3.4) with alternatives starting to appear for the smart phone market on varying operating systems where the input method can vary between a touch screen or stylus input. Touch screens are becoming a part of everyday life, such as point of sales terminals and mobile devices, where the applications may not need high resolution displays with a critical input due to the size of the soft buttons (Sears and Shneiderman, 1991). However, some notational systems may require this critical input. The system presented here used a mouse allowing a far greater level of input when compared to a finger which would take up a larger percentage of the screen, this is where a stylus may offer a better alternative to a finger (Sears and Shneiderman, 1991; Forlines et al., 2007). For these reasons, when notational systems transfer over to a portable/touch screen solution, they may need to be re-tested due to this input change.

6. Conclusion

To measure marksmanship statistics, a valid and reliable system needs to be used. The computerised example shown here allows the input of arrow locations with results showing X and Y raw data coordinates are acceptable and could be improved with experience of the system. Errors are identified mainly when arrows overlap each other, in other target sports, such as pistol shooting where a single hole is left this issue would likely be overlooked or minimised.

Using various statistics to analyse marksmanship allows the coach/athlete/researcher to track varying equipment setups, and changes to biomechanics, physiology and psychology, allowing for continued athlete, sport and equipment development. Previous works in archery could use systems such as this to re-evaluate their results in greater detail than pure score.

7. References

Bannigan, K., Watson, R. (2009). Reliability and validity in a nutshell. **Journal of Clinical Nursing**, 18, 3237–3243

Bar-Eli, M., and Azar, O.H. (2009). Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences. **Soccer and Society**, 10(2), 183-191.

Bowbuddy (Version 3.4) [Computer Software]. UK: T.B.S. Software. Retrieved from: http://bowbuddy.com [Accessed January 2012]

Bruce, L., Farrow, D., Raynor, A., and May, E. (2009). Notation analysis of skill expertise differences in netball. **International Journal of Performance Analysis of Sport**, 9, 245-254.

Choi, H., O'Donoghue, P.G. and Hughes, M.D. (2007), An investigation of inter-operator reliability tests for real-time analysis system. **International Journal of Performance Analysis of Sport**, 7(1), 49-61.

Danage (2012) **Scoring Systems**. Available from: http://www.danagetargets.co.uk/scoringsystems.php [Accessed November 2012]

Di Salvo, V., Collins, A., McNeill, B., Cardinale, M.. (2006), Validation of Prozone ®: A new video-based performance analysis system. **International Journal of Performance Analysis in Sport**, 6(1), 108-119.

Dixon, J.S., and Bird, H.A. (1981). Reproducibility along a 10 cm vertical visual analogue scale. **Annals of Rheumatic Diseases**. 40, 87-89.

Ertan, H., Soylu, A. R., and Korkusuz, F. (2005). Quantification the relationship between FITA scores and EMG skill indexes in archery. **Journal of Electromyography and Kinesiology**, 15, 222–227.

Forlines, C. and Wigdor, D. and Shen, C. and Balakrishnan, R. (2007) Direct-touch vs. mouse input for tabletop displays. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**, 647-656.

Hay, J., & Reid, J. (1988). **Anatomy, Mechanics, and Human Motion**. (2nd Ed..) New Jersey: Prentice Hall.

Hughes, M., and Franks, I.M. (2004). Notational analysis - a review of the literature. In Hughes, M., and Franks, I.M. (Eds.) **Notational Analysis of Sport: Systems for Better Coaching and Performance in Sport**. (pp. 59-106) London: Routledge.

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

IBM (2010). **SPSS** (Version 19) [Statistics Software].

Johnson, R.F. (2001). Statistical Measures of Marksmanship (Publication No. TN-01/2). Retrieved from Military Performance Division website: http://www.dtic.mil/ [Accessed January 2012]

Keast, D., and Elliot, B., (1990), Fine body movements and the cardiac cycle in archery. **Journal of Sports Sciences**, 8(3), 203-213.

Koch, C., and Tilp, M. (2009). Analysis of beach volleyball action sequences of female top athletes. **Journal of Human Sport and Exercise**, 4(3), 272-283.

Leroyer, P., Van Hoecke, J., and Helal, J.N. (1993). Biomechanical study of the final push-pull in archery. **Journal of Sports Sciences**, 11, 63-69.

Lin, K., and Hwang, C. (2005) Analyses of the relationship between the aiming time and the shot points in archery. **Journal of Physical Education in Higher Education**, 7(4), 161-173.

MathWorks, The (2007). **Matlab** (Version 7.5.0.342) [Programming Software].

Maslovat, D., and Franks, I.M. (2007). The need for feedback. In M. Hughes and I.M. Franks (Eds.), **The essentials of performance analysis: An introduction** (pp. 1-7). Abingdon, Oxon: Routledge.

Microsoft (2007). **Excel** (Version 2007) [Spreadsheet Software].

O'Donoghue, P. (2010). **Research methods for sports performance analysis**. Abingdon, Oxon: Routledge.

O'Donoghue, P. (2012). **Statistics for Sport and Exercise Studies**. Abingdon, Oxon: Routledge.

Plumert, J.M. and Kearney, J.K. and Cremer, J.F. and Recker, K.(2005) Distance perception in real and virtual environments. **ACM Transactions on Applied Perception (TAP)**, 2(3), 216-233.

Sears, A. and Shneiderman, B. (1991). High precision touchscreens: design strategies and comparisons with a mouse. **International Journal of Man-Machine Studies**, 34(4), 593-613.

Shrout P.E., and Fleiss, J.L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. **Psychological Bulletin**. 86(2), 420-428.

Van der Kamp, J. (2011), Exploring the merits of perceptual anticipation in the soccer penalty kick. **Motor Control**. 15(3), 342-358.

Vidales, C.E. (2002). How to calibrate touch screens. **Embedded Systems Programming**, 15(6), 32-40.

Vishton, P.M., Rea, J.G., Cutting, J.E. and Nuñez, L.N. (1999) Comparing effects of the horizontal-vertical illusion on grip scaling and judgment: Relative versus absolute, not perception versus action. **Journal of Experimental Psychology: Human Perception and Performance**, 25(6), 1659-1672.

Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. **Journal of Strength and Conditioning Research**, 19(1), 231-240.

Wilkinson, M., Leedale-Brown, D., and Winter, E.M., (2009) Validity of a Squash-Specific Fitness Test. **International Journal of Sports Physiology and Performance**, 4, 29-40.

Wilson, K. and Barnes, CA, (1998) Reliability and validity of a computer based notational analysis system for competitive table tennis. In Lees, A., Maynard, I., Hughes, M. and Reily, T. (Eds.) **Science and Racket Sports II.** (pp. 265-268). London: Routledge.

World Archery (2010a). **FITA Rules and Constitution: BOOK 1 Constitution and Administrative Rules**. Retrieved from: http://www.archery.org [Accessed November 2011]

World Archery (2010b**). FITA Rules and Constitution: BOOK 2 Outdoor Archery Rules**. Retrieved from: http://www.archery.org [Accessed November 2011]