A comparison of visual search strategies of elite and non-elite tennis players through cluster analysis

Nicholas P. Murray & Melissa Hunfalvay

To cite this article: Nicholas P. Murray & Melissa Hunfalvay (2016): A comparison of visual search strategies of elite and non-elite tennis players through cluster analysis, Journal of Sports Sciences, DOI: 10.1080/02640414.2016.1161215

To link to this article: http://dx.doi.org/10.1080/02640414.2016.1161215

Published online: 23 Mar 2016.
A comparison of visual search strategies of elite and non-elite tennis players through cluster analysis

Nicholas P. Murray and Melissa Hunfalvay

Department of Kinesiology, East Carolina University, Greenville, NC, USA; RightEye, LLC, Chief Science Officer, Bethesda, MD, USA

ABSTRACT
Considerable research has documented that successful performance in interceptive tasks (such as return of serve in tennis) is based on the performers’ capability to capture appropriate anticipatory information prior to the flight path of the approaching object. Athletes of higher skill tend to fixate on different locations in the playing environment prior to initiation of a skill than their lesser skilled counterparts. The purpose of this study was to examine visual search behaviour strategies of elite (world ranked) tennis players and non-ranked competitive tennis players (n = 43) utilising cluster analysis. The results of hierarchical (Ward’s method) and nonhierarchical (k means) cluster analyses revealed three different clusters. The clustering method distinguished visual behaviour of high, middle- and low-ranked players. Specifically, high-ranked players demonstrated longer mean fixation duration and lower variation of visual search than middle-and low-ranked players. In conclusion, the results demonstrated that cluster analysis is a useful tool for detecting and analysing the areas of interest for use in experimental analysis of expertise and to distinguish visual search variables among participants.

Sports-specific research has documented that the performance of interceptive actions are an important element in the discrepancy between skill levels across a multitude of sports, including basketball (Oudejans, van de Langenbergh & Hutter, 2002), cricket (Mann, Abernethy, Farrow, Davis & Spratford, 2010), hockey (Panchuk & Vickers, 2006), tennis (Ward, Williams & Bennett, 2002) and volleyball (Lee, 2010). These interceptive actions are the outcome of effective information processing through which the appropriate perceptual information is identified, a response is selected and a motor response is initiated. The perceptual information gathered is used to produce goal-directed actions. Within the perceptual response, the visual system is the dominant sensory system and critical for detecting important relevant cues to effectively produce an appropriate motor response (Zupan & Merfeld, 2005). As such, for successful performance, an athlete, especially in tennis, must identify and attend to relevant visual cues. This process is accomplished through gaze control. Gaze control is an important visual factor that has been found to improve with skill development (e.g., Janelle et al., 2000; Panchuk & Vickers, 2006; Vickers, 1996; Vickers & Lewinski, 2012; Williams, Ward, Knowles & Smeeton, 2002). Gaze control indicates how long and where an athlete visually focuses prior to engaging in a motor skill and is a part of the visual search process. The visual search process is likely context dependent and is influenced by the task, the skill and environmental conditions (Murray & Janelle, 2003).

The conventional approach is to examine expert-novice differences and compare mean differences between these two divergent groups (e.g., Laurent, Ward, Williams & Ripoll, 2006). The typical finding is experts use of few fixation with longer durations (e.g., Perez, Mendez, Manzano & Collado, 2013; Piras, Pierantozzi & Squatrito, 2014) which translate to perceptual advantage that reduces the expert’s processing time and increases response accuracy (Mann, Williams, Ward & Janelle, 2007). In addition, research has also demonstrated that efficient visual search strategy results in improved decision-making through better pattern recognition and the ability to extract critical information from an opponent’s posture (Piras et al., 2014). The measurement of visual search strategies that distinguish experts from novices include differences in fixation location (or the allocation of visual attention to an intended target for 100 ms or longer within 3° of visual angle), search rate (or a measure of the total number of fixations divided by the total fixation duration), visual search order (i.e., scan path) and fixation duration.

Although past research has provided evidence of where experts look compared to novices for a variety of sports, two major limitations for comparing results have been participant selection criteria and the definition of expertise, and the use of averaged data within specific areas of interest. For example, in their seminal work, Goulet, Bard and Fleury (1989) examined visual search patterns of experts and novice tennis players using video format. They found no visual search differences in the preparatory phase but differences in scan paths during the ritual phase between experts and novices. Experts also performed better on identifying serve type (i.e., improved response accuracy). These findings indicated cue usage differences rather than specific visual search strategy differences, as both experts and novices visually selected the same areas of interest.
Research to date has used the term “expert” in tennis for a variety of different skill levels. Tenenbaum et al. (1996) defined experts as professional players with an official Association of Tennis Professionals (ATP) ranking (not limited in range). Singer et al. (1998) defined experts as ranked college level players. Goulet et al. (1989) specified experts as those ranked in the top 40 in Quebec or who were currently or previously ranked players in the Quebec area. Ward et al. (2002) defined experts as club level players with 11.9 years of experience. It is important to provide a very high standard of expert to accurately portray the most appropriate visual search strategies. It is also important that the expert currently holds that status at the time of assessment as the game of tennis is constantly changing due to the advancement of technology and changing techniques and biomechanics.

Important in expertise research is classifying skilled performers who are similar to one another within a group, but meaningfully different from performers of another group. Our approach is to move beyond relying on potentially arbitrary divisions of expertise (such as collegiate, professional or rank) determination by using an empirical, data-driven approach, specifically a cluster analysis (Hair, Black, Babin & Anderson, 2010). Cluster analysis divides data into meaningful and useful groups. It captures the natural structure of the data. The objective is to identify a cluster in which the data is similar (or related) to one another and distinctly different from another cluster. Cluster analysis involves clustering algorithms to compute the distance or similarity matrix between variables and finding a solution that minimises the within-cluster variation and/or maximises the between-cluster variation. The purpose of this study was to examine differences in visual pattern behaviour using cluster analysis to distinguish expertise by the pattern of visual search behaviour in ranked and unranked tennis players. Using a controlled environment, ranked and unranked players will observe a set of video recorded tennis serves to determine if distinct visual search patterns will emerge related to skill. Therefore, it is hypothesised cluster analysis that will differentiate visual search variables, including number of fixations, fixation durations and fixation location during temporal segments of tennis players by skill level and experience. It is also hypothesised that high-ranked players will display fewer fixations of longer duration to areas of the body and longer durations for action phase of the serve.

Method

Participants

A total of 43 tennis players (aged 18–34 years, $M = 23.63$, $SD = 3.59$; 21 males, 22 females) participated in this study. Players were ranked 44th in the world to unranked on the Women’s Tennis Association (WTA) tour or ATP tour for males and had played tennis for an average of 16.74 years ($SD = 4.52$).

Task

Although, some visual search research has demonstrated difference between laboratory based task and naturalistic experimental designs (e.g., Dicks, Button & Davids, 2010), a video model was used to ensure all participants viewed the same tennis serve under the same time constraints and to reduce any potential confounds that may affect our data driven, empirical approach. During viewing the participants’ gaze responses were monitored in terms of number of fixations, fixation durations and fixation location during temporal segments and specific locations throughout the presentation of the serve. A fixation is a period in which the eye is stable for a minimum of 100 ms and does not move more than 1° of visual angle.

Test Film. A professional ranked male player was used as a model and had a current ATP ranking of 32 (age 23 years, 19 years of tennis playing experience). The model was filmed from a “front on” perspective using a digital video camera (Sony, DCR-TRV19). The video camera was positioned at the receiver’s end of the court, 3 ft from where the singles sideline and baseline intersect and towards the hash mark (or centre baseline) on the baseline at a height of 171.45 cm, based on the average height of a male. The model performed 18 serves, 9 from each side of the court. Three types of serves (flat, slice and topspin/kick) were executed and each type of serve was hit in one of three directions (wide, at the body or down the centre) within the service boxes.

The video recording, including sound, was then edited using the Pinnacle Studio Version 7 (Pinnacle Systems, Inc.) editing system. Each serve included four phases of the tennis serve: the first phase, the initiation phase, precedes the initiation of the serve and consists of ball bounces, and foot positioning ($M = 4320$ ms, $SD = 0.93$; Goulet et al., 1989). The second phase, the preparatory phase, begins with the elevation of the arm holding the ball and ends at the apex of the ball trajectory ($M = 900$ ms, $SD = 0.20$; Goulet et al.). The third phase was the execution phase which starts at the servers’ extension up towards ball contact and finishes at ball/racquet contact ($M = 570$ ms, $SD = 0.37$; Goulet et al.). The fourth phase was added (referred to by the researcher as the finishing phase), starting immediately after ball–racquet contact and ending as the ball crossed the net, at which time the video was cut ($M = 280$ ms, $SD = 0.11$). Each serve was shown in real time and included each of the four phases. The average duration of each serve was 6070 ms ($SD = 0.93$). Video editing enabled the 18 serves to be presented in random order and a grey screen was presented for 2 s between each clip.

The Eye-gaze Response Interface Computer Aid (ERICA). An ERICA (2003, model 000–0–103, http://www.eyegaze.com/) system utilised low powered infrared light to track eye movements and interspersed stopping points from the cornea of the eye. These points were then translated into coordinates on the computer screen providing raw data. The data sampling rate was 60 frames/s with an accuracy and precision of ± 0.5” of visual angle. Calibration was automatic using a one-point eye calibration and then a 16-point screen calibration. The videos were displayed on the ERICA system which utilised a 21 inch screen at 60 Hz with a spatial resolution of 0.5”. Participants were positioned 90 cm away using 18.72”x24.28” field of view.

Procedure

Participants were tested at professional tennis tournament sites in various locations and at various times. This study was
approved by the Institutional Review Board and all participants signed an informed consent form. After completing demographic and ranking information, participants were seated in front of the computer and their eyes were calibrated for the ERICA system. They were then said to read a statement of instructions. Participants were asked to watch the serve and "imagine you are on the tennis court playing this person in a competitive match situation, such as at this tournament here in … you are about to return serve during the match… think about and imagine trying to return the serve as effectively as possible making it difficult for the server to return." If the participant had no questions she/he watched three serves presented in random order to familiarise her/him with the video. After the presentation of the three serves if the participant had no further questions, they were checked again for calibration and watched the testing video in its entirety. At the completion of the video, participants filled out a concluding questionnaire to assess the realism they thought the video portrayed. Ninety-five percent of the participants felt the video was either "very realistic" (44%) or "somewhat realistic" (51%).

Data analysis

Visual search variables included the number of fixations (defined as ≥ 100 ms) and fixation durations. These variables were calculated first as the total scores for each serve and then averaged across the eighteen serves. Moving area of interests (AOIs) were established for three different areas: ball, upper body and racquet (Goulet et al., 1989). The moving areas of interest follow the movement pattern of the server and allow for dynamic assessment of visual search characteristics of the participants.

A hierarchical and nonhierarchical cluster analyses were conducted using a two-step process to improve stability in the cluster solution (Hair et al., 2010). Using standardised scores, the observed variables (fixation duration, fixation location difference score and number of fixations) were entered into the cluster analysis. The first stage involved a hierarchical cluster analysis using Ward’s linkage method with squared Euclidian distance measure to determine the number of clusters in the data. Ward’s method is an agglomerative clustering method based on sum-of-squares criterion and produces groups that minimise within-group dispersion (Hair et al., 2010). The second stage involved a k means (nonhierarchical) cluster analysis by specifying the most appropriate cluster solution from Stage 1.

After identifying the visual profiles, we performed separate univariate ANOVAs using the dependent visual control variables (fixation duration, fixation location coordinates (X, Y coordinates of fixation locations) and number of fixations) as well as examined the variance of fixation location to explore differences between clusters. Significant multivariate effects (P < .05) were followed up with post-hoc comparisons between cluster groupings using Bonferroni adjustments as appropriate. Lastly, eta squared (η²) were used to determine effect sizes.

Results

A hierarchical cluster analysis was conducted using Ward’s method with a squared Euclidean distance measure on the standardised visual control variables. The agglomeration schedule coefficient and the dendrogram classified either three or four clusters as two possible solutions. A three-cluster solution was deemed the best fit according to empirical considerations (specific patterns of the observed variables) and how interpretable the cluster solution was. Next, a k means cluster analysis was conducted on the standardised visual control variables for the three-cluster solution. The nonhierarchical solution provided support for the hierarchical analysis. To provide a descriptive indication of the strength of our cluster solution, we conducted a MANOVA on the multivariate effect of the cluster membership. The MANOVA revealed a significant multivariate effect on cluster membership, Wilks’ Lambda = .9994, F(8, 15,290) = 13.476, P < .01, thus indicating reasonable support for our cluster solution (See Figure 1). Figure 1 includes all visual fixation for each cluster layered on top of each other and indicates clear and distinct visual fixation location differences for each cluster.

Descriptive statistics (fixation duration, number of fixations and location variables) for the three clusters (Table 1) were evaluated by separate one-way analysis of variance. The ANOVA for fixation duration was significant, F(2, 7648) = 4.547, P < .1, η² = .321, as well as location variables (x-coordinate, F(2, 7648) = 21.46, P < .01, η² = .468 and y-coordinate, F(2, 7648) = 13.608, P < .01, η² = .318). No cluster differences were found for number of fixations. Tukey post-hoc analysis for fixation duration indicated significant differences between all three clusters (i.e., high, moderate and low ranked demonstrated significantly different fixation duration), whereas cluster three significantly differed from clusters one and two for the location variables (See Figure 2). That is low ranked (Cluster 3; C3) significantly differed from high ranked (Cluster 1; C1) and moderate ranked (Cluster 2; C2) for fixation location. Furthermore, secondary post-hoc analysis using Centre of Bias (the preference to look at the centre of the most informative location of the stimulus) demonstrated that high-ranked players (C1) demonstrated lower variation in fixation movement than did moderate-(C2) and low-ranked (C3) players (p < .05, See Figure 1).

Two separate one-way MANOVA were conducted to compare fixation duration and number of fixations between clusters for each moving area of interest (body, bounce and racquet). Significant differences were found among the three moving area of interests for duration (Wilks’ Lambda = .387, F(6, 74) = 7.212, P < .001, η² = .378) and for number of fixations significant differences were also found (Wilks’ Lambda = .546, F(6, 74) = 4.636, P < .001, η² = .261). The three-cluster solution differentiated cluster membership by rank. High ranked (C1) had mean rank of 60.43 with 88% membership of highest ranked players in the sample. Moderate ranked (C2) demonstrated a mean rank of 206 and 83% membership for moderate-ranked players and low ranked (C3) demonstrated a mean rank of 1005.45 with 81% membership of our sample.

Follow-up separate one-way ANOVAs were conducted to examine differences between the three clusters for each
location. The ANOVA for the number of fixations revealed a significant finding for the body AOI, \( F(2, 39) = 10.411, P < .001, \eta^2 = .348 \) as well as the racquet AOI, \( F(2, 39) = 16.207, P < .05, \eta^2 = .139 \); however, there was not a significant difference on the ball AOI for the number of fixations \( (F(2, 39) = 7.517, P = 3.32, \eta^2 = .055) \) (See Figure 3).

Tukey’s post-hoc test demonstrated that high-ranked players (C1) made fixations to the body than did moderate-(C2) and low-ranked (C3) players. For racquet AOI, low-ranked players (C3) made significantly more fixations on the racquet than high-ranked players (C1) however, moderate players (C2) did not significantly differ from high ranked (C1) or low-ranked (C3) players (see Table 2).

The ANOVA for fixation duration demonstrated a significant finding for the body AOI \( (F(2, 39) = 4.035, P < .05, \eta^2 = .138, \) with Table 1. Standard and raw scores of the visual control variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>High ranked (C1)</th>
<th>Moderate ranked (C2)</th>
<th>Low ranked (C3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (ms)</td>
<td>0.131 (1.04)</td>
<td>0.013 (1.02)</td>
<td>-0.074 (.933)</td>
</tr>
<tr>
<td>Number</td>
<td>-0.034 (.969)</td>
<td>-0.012 (.99)</td>
<td>0.034 (1.019)</td>
</tr>
<tr>
<td>X-location</td>
<td>-0.077 (1.02)</td>
<td>-0.04 (1.04)</td>
<td>0.098 (.886)</td>
</tr>
<tr>
<td>Y-location</td>
<td>0.048 (.91)</td>
<td>0.038 (1.11)</td>
<td>0.0874 (.78)</td>
</tr>
</tbody>
</table>

Figure 1. Fixation data by cluster (Note: Figure 2 represents all the visual fixations within the area of interest separated by cluster).

Note: This figure includes all visual fixations recorded during the serve layered on top. It provides a clear indication of the distinct clusters within the analysis. Eye trackers record fixation location through x and y coordinates as such for the purpose of this figure are considered arbitrary.

Figure 2. Phase differences by cluster.
Discussion

The purpose of this study was to use an empirical, data driven approach to distinguish expertise visual search behaviour. The clustering method presented here represents a robust method to demonstrate distinct difference of visual behaviour that defined three clusters: high, middle and low-ranked players. As a result, the cluster analysis is a useful tool for detecting and analysing the areas of interest for use in experimental analysis of expertise and to distinguish visual search variables among participants. Further differences were found in visual search behaviour between clusters in that high ranked (C1) demonstrated longer mean fixation duration compared to middle (C2) and low ranked (C3). Clusters also differentiated phase differences of the serve timing. During the ritual phase of the serve (preserve), results revealed that high ranked (C1) had longer fixation duration and longer fixations on body areas including shoulder and arm areas. In addition, the high ranked (C1) also demonstrated longer durations for the serve or action phase of the serve when compared to middle-(C2) and low-ranked clusters (C3). The clusters did not differ on postserve fixation durations nor were differences found in ball and racquet locations.

The results of this study are consistent with previous work on return of serve in tennis. Originally, Goulet et al. (1989) found that experts also organised their search around the general body position during the ritual phase. In addition, Ward et al. (2002) using both point light and video displays demonstrated experienced tennis players spent most of their viewing time on the head–shoulder and trunk–hip regions, while novice players spent less time on the trunk–hip region and significantly more viewing time on the racket. There were also fewer fixations while viewing the point-light display in comparison to the conventional film, especially in normal conditions. Both groups had less successive fixations in the head–shoulder and trunk–hip regions while viewing the point-light displays. Shim, Miller and Lutz (2005) more recently compared expert and novice tennis players’ stroke anticipation and found statistical significance of anticipatory accuracy across all skill levels that were greater than chance reactions. Additionally, there was evidence that the participants were able to read the opponents movements, but players would wait until they had sufficient information about the ball’s direction before moving to hit the ball. Overall, participants classified as experts were more accurate in stroke anticipation than those who were categorised as novice level players (Shim et al.).

Overall, these results support the guided search model (Wolfe, 2007) which proposes that basic features direct the viewer’s attention and then selective attention processes are employed. During the selection phase, the most salient features guide the visual search process. The more salient features are more likely fixated; however, if a target lacks saliency then the search process becomes rather random. In addition, the selection of features is proposed to be a serial process. Therefore, distinguishing expert and novice purely on average differences between areas of interest may not tell the whole story. The cluster analysis reveals distinct patterns of visual fixation and indicates an experts visual search patterns have lower variations in movement (as noted in Figure 1). The centre of bias analysis revealed expansion of visual search parameters for moderate-and lower-ranked players. Knowledge about the nature of the scene and the action

Table 2. Number of fixations by cluster.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Body</th>
<th>Ball</th>
<th>Racquet</th>
</tr>
</thead>
<tbody>
<tr>
<td>High ranked (C1)</td>
<td>2.57 (.23)**</td>
<td>1.87 (2.15)</td>
<td>2.45 (.822)*</td>
</tr>
<tr>
<td>Moderate ranked (C2)</td>
<td>4.25 (1.15)**</td>
<td>2.52 (1.41)</td>
<td>3.96 (2.57)</td>
</tr>
<tr>
<td>Low ranked (C3)</td>
<td>5.27 (2.14)**</td>
<td>2.93 (1.90)</td>
<td>4.63 (2.90)*</td>
</tr>
</tbody>
</table>

* P < .05; ** P < .001

Table 3. Fixation duration by cluster.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Body</th>
<th>Ball</th>
<th>Racquet</th>
</tr>
</thead>
<tbody>
<tr>
<td>High ranked (C1)</td>
<td>0.952 (.0332)*</td>
<td>0.317 (.0261)**</td>
<td>0.903 (.0601)**</td>
</tr>
<tr>
<td>Moderate ranked (C2)</td>
<td>0.423 (.0364)</td>
<td>0.425 (.0339)</td>
<td>0.859 (.0750)**</td>
</tr>
<tr>
<td>Low ranked (C3)</td>
<td>0.361 (.0309)*</td>
<td>0.525 (.0287)</td>
<td>0.574 (.0407)**</td>
</tr>
</tbody>
</table>

*P < .05; **P < .001

Figure 3. Fixation duration and location by cluster.
required will guide search in a top–down manner. A skilful person will orient visual attention based on their semantic knowledge of the sport, often within the very first eye movement (Castelhano & Henderson, 2007).

We purposefully conducted this project within a controlled laboratory environment to test the utility of the cluster analysis. This is a potential limitation especially considering recent work that has demonstrated differences between film and in situ conditions (Afonso, Garganta, McRobert, Williams & Mesquita, 2014). It is important to note that in situ often differ in amplitude, but not in direction. For example, Afonso, et al, examining volleyball players, found participants displayed longer fixations for in situ conditions than in film-based conditions. However, there were no differences between fixation location and gaze fixations. Although it may be recommended to use in situ condition when possible, film-based work still provides an opportunity to examine visual search behaviour expertise and skilled decision-making in sport.

This project confirms distinctions between players by rank and expertise level and extends the use of visual behaviour to distinguish expertise level. However, while most research has demonstrated skilled players use a more efficient search pattern (i.e., longer fixation duration with fewer fixations), this is not without controversy as search patterns tend to vary across sport, task and likely the situation (Afonso et al., 2014). Using a more elaborate statistical technique such as a cluster analysis might shed further light on to specific patterning of visual strategies. That is future research should include cluster analysis for area of interest rather than a priori decisions about areas of interest (e.g., predetermined areas of interest – racquet, body, ball, etc…). That is to use visual behaviour to determine the meaningful areas of the scene. Theoretically, groups should cluster by location and duration, therefore, further providing investigations into higher-level questions about the meaning of fixation locations. Granted this approach is a data-driven and descriptive, however, it is potentially a more robust method to examine visual behaviour as algorithms could be utilised to reduce outliers or to parse data by clusters in the context of the visual scene.

Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Nicholas P. Murray http://orcid.org/0000-0001-9968-2109

References


