

Perceptual Skill Identification in a Complex Sport Setting

Ralf Kredel* Andre Klostermann* Olivia Lienhard* Johan Koedijker* Katja Michel*
Ernst-Joachim Hossner*
(*University of Bern, Switzerland
E-mail: ralf.kredel@ispw.unibe.ch

Abstract

Decision making in sports has been extensively studied over the last two decades, with a particular focus on differences in visual search strategies and movement initiation times between experts and novices. The current study expands this paradigm by identifying situation, gender, and expertise specific gaze strategies in performing defensive actions in beach volleyball. Sixty-four beach volleyball players were confronted with 96 scenes displaying 3 attacking variations. The experimental set-up allowed participants to react as they would on court, while decision accuracy and movement initiation time were measured using a 10-camera-VICON-system. Furthermore, gaze behavior was recorded pursuing a novel integrated and automated approach with a high resolution eye tracker (EyeSeeCam, 220 Hz). First analyses show that elite players differ from near-elite players by having a higher percentage of correct decisions combined with a later movement initiation time. Apparently, elite players optimize the time available for information pick-up before initiating their defensive action.

1. Introduction

There is a vast amount of research trying to identify perceptual-cognitive mediators of expert performance and decision making in sport tasks (for a review see [1]). In this scope, decision making seems highly dependent on visual information. Therefore, most research had concentrated on the manipulation of visual stimuli, presented as images or videos using spatial or temporal occlusion techniques. Alternatively, gaze behavior was analyzed using eye tracking equipment, for example in [2].

Thus far, only a few studies, e.g. [3], analyzed the decision making process in real world settings by using mobile eye tracking equipment. Due to the large amount of manual data analysis involved in this method, most studies comprise of a very small number

of participants. Furthermore, most mobile eye tracking devices offer only a low temporal resolution, so that a proper analysis of gaze kinematics (e.g. saccadic movements) is nearly impossible. Additionally, own tests with different mobile eye tracking devices revealed that in dynamic situations the accuracy of the results obtained is fairly low and the measured gaze positions need to be continuously readjusted by “educated guess” of the current position. Last, most mobile eye tracking devices only allow for very small mobility as participants get directly connected to heavy recording devices, which need to be powered, or need to stay close to a desktop, as the length of cabling is limited.

Thus, the question arises whether the inferences mentioned above provide sufficient validity for generalization, especially in the research on expertise. In particular, as it has been shown in [4], that experts’ perceptual advantages increase with higher “in situ”-ness of the response, while most laboratory studies allowed for spatially limited responses (verbal, button press, contact-mats) only.

Consequently, from a theoretical as well as an applied point of view, it seems necessary to identify gaze strategies by equipment with high temporal and spatial resolution, while keeping both, stimuli as well as responses, similar to the real-world situation. Therefore, we integrated a lightweight, mobile eye tracking device (EyeSeeCam, [5]) into our sensorimotor lab, allowing for synchronous recording of gross-scale motion with the participant’s gaze in the laboratory frame-of-reference, while displaying life-size visual stimuli. In the following, results of the first study will be described with a special emphasis on the methodology involved.

2. Design

Using a mixed factorial design, the defense behavior of 64 beach-volleyball players of the Swiss

national league was tested (32 female / 32 male; 32 elite / 32 near-elite).

Participants watched six videos (3 near-elite and 3 elite models) of 20 attacking scenes on a 3x4 m screen from the perspective of the defenders' position in the court's left half (3m away from the screen with a slight offset to the left, see Fig. 1). After 4 warm-up scenes, each video contained 8 smash, 4 cut and 4 line shots in randomized order.



Fig. 1: Example scene of the stimulus material (perspective of the defender).

Participants were instructed to react to the type of attack as they would do on the field. A long line shot would result in a movement to the right rear corner of the field, while the correct solution for a cut would be a movement to the front left. After a smash, participants had to hold their position indicating a defense action with their arms only (= no-go trials).

As dependent variables, decision correctness, movement initiation time as well as number of fixations and total fixation duration were calculated so far, while the body and gaze kinematics for the whole trial are available for further analysis.

3. Methods

3.1 Eye Tracker EyeSeeCam

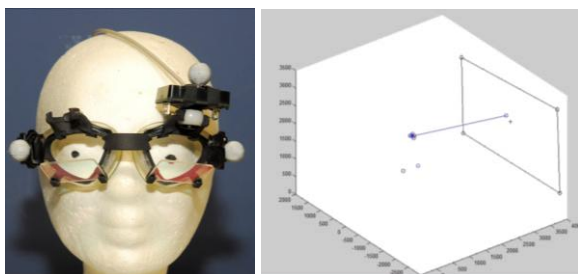


Fig. 2: Left: reduced EyeSeeCam version with retro-reflective markers used to track head position and orientation. Right: 3D lab space with gaze vector, screen and additional markers (hand, head, etc.).

The basis of the eye tracking system is a reduced version of the EyeSeeCam presented in [5]. Instead of having a scene camera and calibration lasers, the goggles consist only of a single IR camera, mounted directly to the frame, which renders the system very lightweight (Fig. 2, left panel).

Additionally, participants were wearing a small bum bag with an optical bridge, needed for the extension of the Firewire connection to the control laptop, and its power supply.

The EyeSeeCam software on the control laptop performs image processing on the obtained video stream and the extraction of the pupil's center as well as the corneal reflections to provide horizontal and vertical eye rotation angles in real-time. Additionally, the calculated angles are mapped to two channels of an analog output card as a voltage signal. The overall latency from image acquisition to analog output depends on the adjusted frame rate of the camera ([5]) and can be calculated as:

$$\text{latency} = 1 \text{ frame period} + 0.5\text{ms}$$

In our setting (220Hz), the resulting latency is 5ms (which is in line with the typical delay of the Vicon system at 200Hz).

The analog card is connected to the A/D interface of the Vicon system, which is synchronously sampled with respect to Vicon camera acquisition (with variable oversampling). As the Vicon system provides the possibility to access kinematic marker data and analog data in real-time via Ethernet stream, eye rotation angles as well as head position and orientation is now accessible in real-time and can be processed further, e. g., by a Matlab script (Fig. 2, right panel) or by VR Juggler, a virtual reality framework with VRPN support, [6].

3.2 Integration Software

The eye tracking software requires a calibration procedure to output accurate rotation angles of the eye. Typically, the EyeSeeCam tracker has a laser LED attached, which projects a pattern of five dots in a regular grid with a distance of 8.5° between the dots onto a wall or screen. As this projection is already head-centric, the angular displacement of the eyes while fixating these dots correspond to the angular position of the dots ($0^\circ/0^\circ$, $-8.5^\circ/0$, $8.5^\circ/0$, $0^\circ/-8.5^\circ$, $0^\circ/8.5^\circ$). The corresponding rotational motion of the eye in the video stream (in pixel) can then be mapped to angles (in degree). As the reduced version of the eye tracker does not have this laser grid installed, the projected dots need to be emulated on a screen.

In order to do so, we developed a Matlab script, which calculates the positions of the five dots on a

screen with known dimensions and position relative to the lab reference frame. The dots' positions are constantly updated based on the current head position \mathbf{t}_{head} and orientation \mathbf{R}_{head} of the participant with an update rate of 25Hz (what is not crucial as the calibration is done in a quasi-static condition of the head). A further rotational adjustment \mathbf{R}_{adj} was integrated, that allowed for the adjustment of the “zero viewing position”, which is different for each person due to individual physiognomy and the resulting position of the goggles relative to the eyes. In sum, the coordinates of eye position \mathbf{t}_{eye} and gaze direction \mathbf{t}_{gaze} can then be calculated as:

$$\mathbf{t}_{\text{eye}} = (\mathbf{R}_{\text{head}} * (\mathbf{R}_{\text{adj}} * \mathbf{t}_{\text{eyeOffset}})) + \mathbf{t}_{\text{head}}$$

$$\mathbf{t}_{\text{gaze}} = (\mathbf{R}_{\text{head}} * (\mathbf{R}_{\text{adj}} * ((\mathbf{R}_{\text{eye}} * [\lambda \ 0 \ 0]^T) + \mathbf{t}_{\text{eyeOffset}}))) + \mathbf{t}_{\text{head}}$$

with $\mathbf{t}_{\text{eyeOffset}}$ being the distance of the center of the eye ball to the origin of the head-local coordinate system and λ an arbitrary scale factor for the length of the gaze vector. During the calibration procedure, this participant-specific information (\mathbf{R}_{adj} , $\mathbf{t}_{\text{eyeOffset}}$) is stored to disk for later off-line analyses.

By calculating the intersection of the gaze vector with the screen plane, the current gaze point on the screen can be identified online. This feature is used to verify the accuracy of the calibration process by calculating the RMS (root mean square) between projected points and gaze position, while instructing the participant to fixate the dots.

After system calibration, the gaze vector can be displayed in the lab reference frame, which opens new possibilities compared to head centered gaze vector calculations. By this means, not only head/eye coordination is measurable, but also gaze interactions with other predefined objects in the 3D lab space. By digitizing displayed video stimuli or by using VR techniques, gaze position can be algorithmically mapped onto objects in the stimulus material, e. g., by evaluating the objects in different physiological areas of the visual field (fovea centralis, peri- and parafovea). This removes some fundamental restrictions to eye tracking mentioned in [7] and hereby broadens its application possibilities, e. g., with respect to peripheral information pickup. Furthermore, displayed stimuli can be adapted online by the calculated gaze position.

3.3 Decision Time and Direction

The participants' motion was tracked using a 10-camera-Vicon-T20-system. In order to reliably measure the instant of movement initiation and the direction of motion, we used a full body marker model

to calculate a “kinematic body center” by the mean of the marker trajectories.

The projected motion of this body center onto the x/y ground plane is further used to identify decision time and direction of movement. While watching the scenes, the participant was free to choose his/her preferred position. An individual zero position was calculated (movement of the body center smaller 0.25m/s over 500ms as close to hand/ball contact as possible). For identifying movement initiation, the body center had to leave a virtual circle of 750mm diameter around this zero position. An algorithm then performed a backwards search from this time on until the velocity of the body center was smaller than 0.25m/s. This moment, related to the hand/ball contact of the attacker, was used for the definition of “decision time”. At the moment of leaving the circle the direction vector was mapped onto the main axes of decision (cut: 50°, long line: 270°, with 0° pointing in direction of the screen) and the absolute value was taken to distinguish between cut and long line reactions. As smashes were designed as no-go trials, these trials were not taken into further account.

4. Results and Discussion

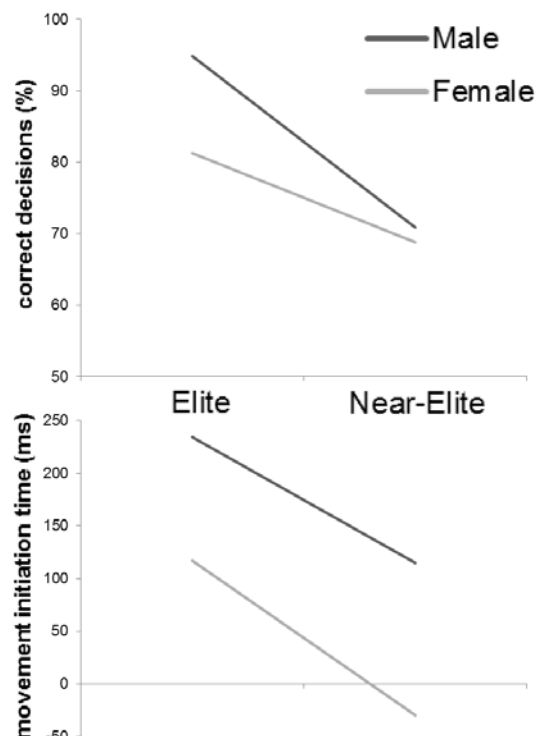


Fig. 3: Correct decisions (%) and movement initiation time (ms) for elite (male/female) and near-elite (male/female) players.

First behavioral results (Fig. 3) show, that elite players made significantly better decisions, $F(1, 23) = 4.88, p < .05, \eta^2 = .20$, than near-elite players. Also, movement initiation times were later for elite players compared to near-elite players, $F(1, 23) = 7.04, p < .05, \eta^2 = .27$. Furthermore, independent of players' skill level, males initiated the movement significantly later than females, $F(1, 23) = 5.41, p < .05, \eta^2 = .22$.

On the one hand, results show again that in situations in which a decision has to be made experts are able to extract more relevant information to reach a higher percentage of correct answers compared to near-elite players. Interestingly, and in contrast to previous reports ([1]), experts, especially males, make use of their physical advantage and wait longer for maximizing information pick-up.

First algorithmic analyses of the players' gaze strategies (Fig. 4) reveal a similar pattern. Both, elite players, $F(1, 23) = 7.84, p < .05, \eta^2 = .29$, and males, $F(1, 23) = 9.81, p < .05, \eta^2 = .34$, had a significant longer total information pick-up than near-elite players and females. Additionally, an interaction between expertise and gender just missed significance level, $F(1, 23) = 7.84, p = .79, \eta^2 = .15$. Interestingly, males (independent of skill) had a higher number of fixations, $F(1, 23) = 12.82, p < .01, \eta^2 = .40$. It may be that due to their later movement initiation males had the chance to take one more cue into account.

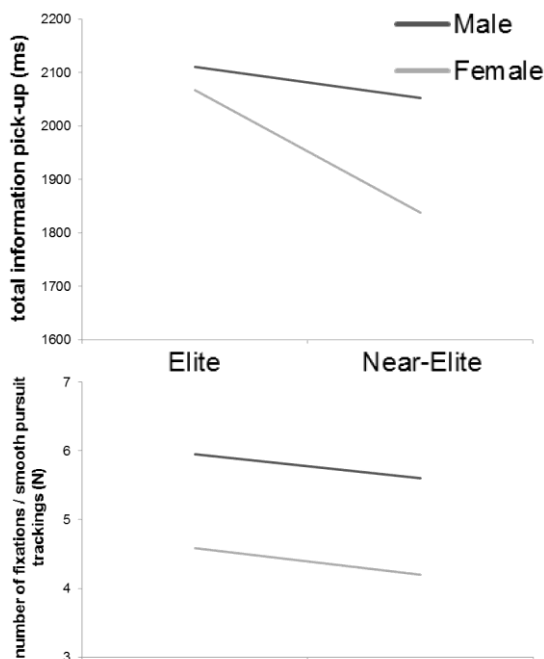


Fig. 4: Total information pick-up (ms) and number of fixations / smooth pursuit trackings (N) for elite (male/female) and near-elite (male/female) players.

Further analysis on the gaze data is carried out at the moment and will bring more insights on those questions, particularly combining cues and gaze patterns. After a successful identification of the most beneficial gaze patterns, the system's online capabilities predestine it to evaluate perceptual skill training possibilities in complex sport settings, e.g., in beach volleyball.

References

- [1] Mann, D.T.Y, Williams, A.M, Ward, P., & Janelle, C.M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport & Exercise Psychology*, 29, 457-478.
- [2] Farrow, D. & Abernethy, B. (2007). Wahrnehmung von Expertinnen und Experten im Sport: Einige Kernfragen und -probleme. In N. Hagemann, M. Tietjens & B. Strauß (Eds.), *Psychologie der sportlichen Höchstleistung* (pp. 71-92). Göttingen: Hogrefe.
- [3] Wood, G. & Wilson, M. R. (2010). Gaze behavior and shooting strategies in football penalty kicks: Implications of a 'keeper-dependent' approach. *International Journal of Sport Psychology*, 41, 293-312.
- [4] David L. Mann, Bruce Abernethy & Damian Farrow (2010). Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks, *Acta Psychologica*, 135, 1, 17-23.
- [5] Kohlbecher, S., Bartl, K., Bardins, S. & Schneider, E. (2010). Low-latency combined eye and head tracking system for teleoperating a robotic head in real-time. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (ETRA '10)*. ACM, New York, NY, USA, 117-120.
- [6] Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A., and Cruz-Neira, C. (2008). VR juggler: a virtual platform for virtual reality application development. In *ACM SIGGRAPH ASIA 2008 courses (SIGGRAPH Asia '08)*. ACM, New York, NY, USA.
- [7] Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. In: *Research Quarterly for Exercise and Sport*, 69, 111-128.