



# The effect of eye movement sonification on visual search patterns and anticipation in novices

Maryam Khalaji<sup>1,2</sup> · Mahin Aghdaei<sup>2</sup> · Alireza Farsi<sup>2</sup> · Alessandro Piras<sup>3</sup>

Received: 4 January 2021 / Accepted: 7 September 2021 / Published online: 18 October 2021  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

## Abstract

Visual information is essential to successfully anticipate the direction of the shot in ball sports whereas using another sense in motor learning has received less attention. The present study aimed to examine whether the multisensory learning with the orienting visual attention through the sound would influence anticipatory judgments with respect to the visual system alone. Forty novice students were randomly divided into visual and audio-visual groups. The experimental procedure comprised two phases; 1 training and 3 testing phases, respectively. During the training sessions, 200 video clips were employed to anticipate the direction of the shot, interspersed by 5-min of rest every 25 trials. A sound was used to orient the attention of the audio-visual group toward key points meanwhile the visual group watched the videos without sounds. Then, during the testing phases, they watched 20 video clips in the pretest, immediate retention, and delay retention test. The film was occluded at the racket-ball contact and then they quickly and carefully decided the direction of the shot. The audio-visual group showed higher response accuracy and shorter decision time than the visual group in the immediate and delayed retention. The audio-visual group exhibited longer fixation duration to the key areas than the visual group. In conclusion, using multisensory learning may not only reallocate perceptual and cognitive workload but also could reduce distraction, since, unlike visual perception, auditory perception requires neither specific athlete orientation nor a focus of attention. In general, the use of the multisensory learning is likely to be effective in learning complex motor tasks, facilitating the discovery of the new task needs and helping to perceive the exercise structure.

**Keywords** Temporal occlusion · Multi-sensory learning · Badminton · Eye movement sonification · Anticipation

## 1 Introduction

Athletes' successful performance is determined by quick and effective responding capabilities where uncertainty is present [1, 2]. Fast movements of hand and foot, as well as the skill to change directions immediately have been assumed to be the sole ingredients of quickness in sports [3]. Moreover, cognitive factors are directed immediately before and during any movement can greatly affect the quality and quickness

of the performance [4–6]. A key cognitive element in elite sport is the ability to anticipate opponent actions before their action completion, such as racket-ball or foot-to-ball contact [2, 3]. Moreover, some researchers have established that athletes change the areas to fixate across the ongoing movement, since the amount or type of kinematic information emanating from different opponent's body areas changes across the movement [5, 7, 8]. For example, Prigent et al., (2014) have investigated how observers estimated a ball's landing point thrown by a virtual agent with different information from its body language. They found that body movements and facial expressions carry-on information about opponent's effort [9]. In the other study, Triolet et al. (2012) proposed that early anticipation behaviors occurred when players use significant context-specific information before the opponent's stroke. When such information are not available, players produce anticipation behaviors at the racket-ball contact using information extracted from the opponent's preparation of the stroke [10]. In the other study Sáenz-Moncaleano, Base-

✉ Mahin Aghdaei  
a-aghdaei@sbu.ac.ir

Alireza Farsi  
a\_farsi@sbu.ac.ir

<sup>1</sup> Shahid Beheshti University, Tehran, Iran

<sup>2</sup> Department of Motor Behavior, Shahid Beheshti University, Tehran, Iran

<sup>3</sup> Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy

vitch, and Tenenbaum (2018) investigated gaze behaviors in high- and intermediate-skill tennis players while they performed tennis serve returns. Results revealed that high-skilled players exhibited better return shots than their lower skill counterparts, showing longer fixation durations on the ball at the pre-bounce [11, 12].

Based on different studies, novices focused more on distal cues (i.e., opponent's head) to predict the location of the ball in comparison to the expert player who fixated on more proximal cues (e.g., trunk, arms) [6, 7, 13]. Practically, we all know that it is not possible to bring the perceptual performance (and anticipatory skills) of a novice player to a professional level simply by forcing them to adopt an expert's perceptual strategy i.e., in the case of badminton, by instructing the novice players to foveate more on the opposing player's proximal body part [8, 13]. As proposed by Papin (1984), the modeling of the expert's perceptual strategy per se is unlikely to be a successful means of enhancing perceptual performance. The relationship between different elements of the body, necessary for predicting subsequent action, is learnable, and some facility is afforded for concomitant development of the knowledge base upon which the expert's perceptual strategy is based [14]. Moreover, orienting attention, covert and overt [12], toward a specific area or a certain object also makes possible to recognize a stimulus more rapidly and identify it more precisely [4, 7, 15, 16].

In the recent years, studies conducted highlight the growing importance of multisensory learning as a component of cognitive learning and a way of better deal with different environmental conditions. Thus, effective multisensory learning strategies accelerate the learning process in various educational areas, including the sports discipline. Studies on learning have emphasized that, when involving more sensory organs, the retention and recall of information is stronger [17]. In this way, different studies have found that using sound along with the visual sense influence the orienting of attention [16, 18–22]. Furthermore, learning perceptual skills is often slow and effortful in adults, and studies of perceptual learning show that a simple task of detecting a subtle visual stimulus can require a month or more of training [23]. As with multisensory learning, novices can practice with a developed and more responsive perceptual-motor workspace, which uses sound as a helpful constraint on the action and its effects on visual perception and visual learning [18, 20, 22].

A plethora of studies have revealed the benefits of multisensory learning over unisensory [16, 22, 24]. The intention of using auditory feedback in interactive systems has recently received higher attention in different research areas. In sport situations, the use of auditory feedback can accomplish with visual feedback. This process, also called sonification, can be used to enhance the performance of the human perceptual system in both motor control and motor learning. Coutrot and colleagues (2012) investigated the effect of the soundtrack

on eye movements throughout the video task. The results revealed that sound might influence fixation location, fixation period, and saccade amplitude [24]. Losing et al., (2014) employed gaze-contingent auditory feedback (sonification) to direct visual attention. They showed that participants could use this direction to speed up visual search tasks [16]. In the other study Canal-Bruland et al., (2018) investigated the effect of auditory contributions to visual anticipation in tennis. The results highlighted that the louder is the sound of the racquet-to-ball contact, the longer is the pursuit of the ball's trajectory [25]. However, sport circumstances are defined by a large amount of complicated visual information, necessitating to recognize and choose the most informative and relevant signals. Predominantly, in racket sports, in which the receiver has to react to the opponent's shot, focusing attention to appropriate body areas has shown to be the key factor for professional execution [26].

During the last decade, a growing number of applications involving the auditory modality to inform or guide users has been developed in a large number of domains, ranging from guidance for visually impaired, positional guidance in surgery, rehabilitation for patients with disabilities, learning of handwriting movement, or gesture guidance [27]. These applications are based on the sonification concept which is defined as “*the use of non-speech audio to convey information or perceptual data*” Despite an increasing number of studies based on sonification for auditory guidance, there is a lack of fundamental studies aiming at evaluating the efficiency of specific sound characteristics to guide user's eye movement towards specific cues of the opponent's body that are important for action perception and anticipation. The present study aims to guide participant's visual search strategy to relevant cues necessary for accurate and quick anticipation of shot direction with the eye movement sonification.

## 2 Materials and methods

### 2.1 Participants

Five expert badminton players (2 females, 3 males, mean age 25.85 years, SD 1.20 years) and forty novices (20 females, 20 males, mean age 24.01 years, SD 3.05 years) were recruited for the study. Novices had not participated in any sport at a professional level, and although they all knew the rules and the practice of badminton, they had never participated regularly in this sport. They voluntarily underwent the test, which did not include any invasive or harmful procedures. The subjects were healthy, had normal hearing and normal vision and they had no prior experience with the task. All participants received a verbal explanation of experimental procedures and gave their written informed consent before participating. The Institutional Ethics Committee of

the Shahid Beheshti University approved the experimental protocol (IR.SBU.ICBS.98/1002).

## 2.2 Apparatus and procedure

### 2.2.1 Test film

The setup was the same used in previous research [7]. A total of 20 video clips of badminton clear forehand, performed by five national league players, were filmed from defense player's perspective. The filming perspective provided a wide viewing angle, necessary to facilitate the depth perception. During the recording sequence, the player was positioned in the middle of the badminton court, hitting the ball and directing it toward the left or the right of the participant's point of view. Clear forehand in badminton was recorded with a video camera and digitalized in a format for anticipation. The videos were integrated into a specially written software program. The program presents the video clip with a resolution of  $1280 \times 960$  pixels in the upper right corner of a display (see Fig. 1).

### 2.2.2 Visual behavior sonification

Five national league players were asked to watch 20 video clips dressing the eye-tracking device for the pre-study results, necessary to determine the best visual search behavior to predict the shot direction in badminton. We used both temporal and spatial occlusion techniques. Various video sequences were produced in which one region, the trunk, was masked with an opaque patch for a fixed time interval. If an important visual cue is occluded during that time interval, this should result in a decline in anticipatory performance [8, 28]. With the temporal occlusion technique (4 levels, from 160 ms before racket-shuttle contact until 80 ms

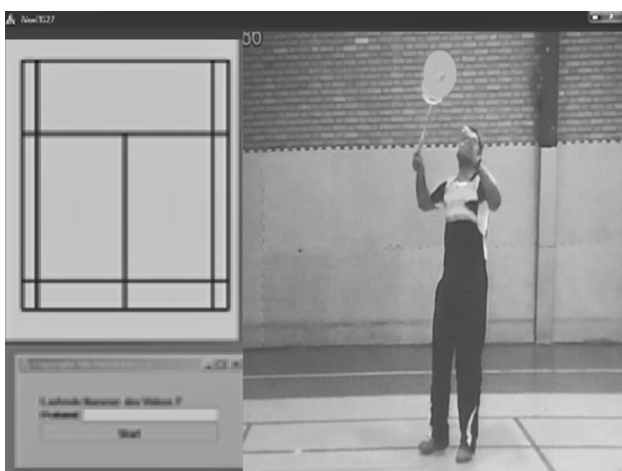
after contact), we can infer which visual cue is important and which is the time interval necessary to predict the badminton shot. We defined 5 interest areas (trunk, legs, hand, racket, and head) in accordance to Hagemann et al. (2006), without considering the shoulders and the background areas. Expert's eye movement data were analyzed, and the visual search pattern was exported from the iView ETG software (trunk up to 160 ms before racket-shuttle contact, the arm region from 160 to 80 ms, and the racket region from 80 ms to shuttle contact) and algorithmically transformed and sonified using parameter mapping sonification, in which the amount of data and its distances from the reference in time scales were mapped to acoustic parameters (reference point sonification) [7]. The sonification was generated by the (C++) software and displayed through standard stereo headphones (frequency range: 14 Hz to 26 kHz) and the broadcast sound was adjusted according to each participant's hearing range.

### 2.2.3 Task

Novice badminton players were randomly assigned to two groups; the visual and audio-visual groups, in which 20 participants predicted the direction of the shuttle only with the visual system, whereas the other 20 participants with the audio-visual system.

Participants became familiar with the procedure through a brief explanation about the method, auditory parameter, and sonification. After introducing the method, interventions and evaluations were carried out at the Physical Education Laboratory of Shahid Beheshti University. Participants sat on a chair in front of a vertical translucent screen, 150 cm from the subject's eyes. The sequences were back-projected by a digital projector, distant 250 cm from the screen, forming an image of 87 cm high and 120 cm wide [13]. Eye movements were recorded binocularly with an eye-tracking system (SMI Eye Tracking Glasses, gaze sampling rates from 50 Hz up to 200 Hz, ETG 2w, Germany). We then calibrated the eye-tracking system with 3-point calibration, 50HZ sampling rate, and verified its functionality by asking the participants to look at a fixed point on the screen to check for accuracy. Each video clip started with a 500 ms acoustic tone to prepare participants for the video onset. The duration of each clip (4 s) was equal in all trials. The film consisted of a preparation phase and racket-ball contact phase. The film was interrupted (temporal occlusion) at the racket-ball contact. After the video clip was occluded, participants had to make a quick decision on whether the shot was left or right pressing the mouse button.

The direction of the badminton shots was completely randomized but kept in the same order for each participant. Twenty video clips of 4 s, with shots directed to the left and to the right were shown to each participant with ten repetitions each.



**Fig. 1** A screen of badminton shot

### 2.2.4 Training

The visual group received verbal feedback in 50% of 200 trials and were given a verbal explanation in which areas they should look at to anticipate correctly the direction of the shot. When participants sat on the chair and film got ready for showing, one of the authors, who was a badminton player and coach, told them “*when the video clip started, please look at the trunk at first, then hand and then racket for anticipation and after that anticipate the direction of the shot*”. Participants received a feedback with their anticipation result as soon as they pressed the mouse button in odd trials. Instead, for the other 50% of trials, they did not get any feedback. This group used verbal feedback and their eye movements were recorded.

The audio-visual subjects wore headphones and were asked to synchronize their eye-movement sound with a sonified reference visual search’s pattern (trunk 160 ms before racket-shuttle contact, the arm region from 160 to 80 ms, and the racket region from 80 ms to shuttle contact) in 50% of 200 trials. For this, sonification of movement error was used. The current deviation between eye movement of participants and optimal visual search was acoustically presented [16, 29]. Indeed, this group used guided exploratory learning method to learn key points for anticipation while they were told before the experiment started in what situation a pleasant and unpleasant sound they would hear. They were told “*when eye movements will move away from the reference points, you will hear an unpleasant sound, and when they will approach the key points, the pleasant sound will be played*”. This group used sound to predict the direction of the shot in odd trials, whereas during the even trials they anticipated the direction of the shot without any feedback.

### 2.2.5 Testing

There was a 5-min break between 25 trials of the experiment. Depending on individual decision time and response speed, the experiment lasted about 70 min to complete for each participant. The testing procedure consisted of 3 phases: pretest, immediate retention, and delayed retention.

*Pretest* refers to the commonly employed evaluation by testing what occurs at the beginning of the test before the experiment was started. The results of this test were considered as a baseline and was done a day before the training sessions. Participants performed 200 trials.

*Immediate retention* the level of performance on this test reveal the gain or loss in the capability for performance in one task as a result of practice or experience on the same task, at least 2-h after the last training session. It was taken 2-h after intervention.

*Delayed retention* was done 7 days after the last training session to measure retained knowledge. Participants were

tested in just visual practice without sound (because any sound is not heard in the real film).

After the study, all participants completed a short questionnaire about which type of training they preferred. The questionnaire was given to the subjects after the delayed retention test to assess three aspects of training. Firstly, they could rate the comfort to train with visual or audio-visual training; secondly, how useful it was for learning the task; and thirdly, how active they were in using the training. The answer scale ranged from 1 (not at all) to 7 (very much) [29, 30].

## 2.3 Dependent variables and analysis

### 2.3.1 Anticipation test

The ability to make accurate predictions from advanced sources of information was measured as follows.

### 2.3.2 Response accuracy (RA)

The percentage of trials in which the subject’s response was correct or incorrect was determined.

### 2.3.3 Decision time (DT)

The time (ms) from racket-ball contact to the button pressed by the participant was determined. Responses with RTs shorter than 150 ms and longer than 600 ms were discarded (early or delayed responses).

### 2.3.4 Visual search data

The following visual search measures were analyzed.

### 2.3.5 Search rate

This measure included the mean number of fixation locations per trial and the mean fixation duration (in milliseconds) per trial.

Response accuracy, decision time, and search rate were analyzed separately using a 2 (group: visual and audio-visual) \* 3 (time: pretest, immediate retention, and delay retention) repeated-measures ANOVA to assess interaction and between-group effects. If the interaction was significant, follow-up repeated-measures ANOVA was also used to assess learning within each time. The post hoc Bonferroni test was applied for multiple comparisons between the tests. Violations of sphericity were corrected using Greenhouse–Geisser correction. An independent t-test was also used to assess whether the groups differed at different times at the learning rate.



### 2.3.6 Viewing time in different location

The mean fixation duration demonstrated that participants spent fixating the gaze on each interest area of the display when trying to anticipate the shuttle direction was determined. The display was divided into five fixation locations: head, trunk, feet, hand, and racket areas. Fixation duration was analyzed using a factorial analysis of variance (ANOVA) in which group was a between-participant factor, and fixation location was a within-participant factor.

## 3 Result

### 3.1 Response accuracy

The result of repeated measure ANOVA revealed significant interaction between group and time ( $F_{(2, 76)} = 502.16$ ,  $p = 0.001$ ,  $\eta^2 = 0.93$ ). After that, a repeated measure ANOVA and independent sample t-test was also done to analyze response accuracy in each time and group separately.

Repeated measure ANOVA test indicated a significant main effect for time ( $F_{(2, 38)} = 347.49$ ,  $p = 0.001$ ,  $\eta^2 = 0.95$ ) and independent sample t-test revealed significant difference between groups in immediate retention ( $t = -3.74$ ,  $df = 100$ ,  $p = 0.001$ ) and delayed retention ( $t = -3.36$ ,  $df = 100$ ,  $p = 0.001$ ).

The result of the Bonferroni correction test revealed that there is a significant difference between the pretest and immediate retention in audio-visual group ( $p < 0.05$ ) and response accuracy in immediate retention ( $72.22 \pm 10.01$ ) was higher than pretest ( $38.10 \pm 8.72$ ). Also, there is significant difference between pretest and delayed retention ( $p < 0.05$ ) and response accuracy in delayed retention ( $65.96 \pm 13.96$ ) was higher than pretest ( $38.10 \pm 8.72$ ) but between delayed retention and immediate retention, there is no significant difference.

In visual group, the result of the Bonferroni correction test revealed that there is a significant difference between the pretest and immediate retention ( $p < 0.05$ ) and response accuracy in immediate retention ( $57.38 \pm 13.52$ ) was higher than pretest ( $39.83 \pm 9.39$ ). Also, there is significant difference between pretest and delayed retention ( $p < 0.05$ ) and response accuracy in delayed retention ( $51.77 \pm 15.74$ ) was higher than pretest ( $39.83 \pm 9.39$ ) and between delayed retention and immediate retention, there is a significant difference ( $p < 0.05$ ). Response accuracy in immediate retention ( $57.38 \pm 13.52$ ) was higher than delayed retention ( $51.77 \pm 15.74$ ) (Fig. 2).

### 3.2 Decision time

A repeated measure ANOVA ( $2 \times 3$ ) revealed that there is a significant interaction between group and time ( $F_{(2, 76)} = 566.91$ ,  $p = 0.001$ ,  $\eta^2 = 0.94$ ). After investigating interaction, ANOVA showed a significant effect for the time in the audio-visual group ( $F_{(2, 38)} = 1225.21$ ,  $P = 0.001$ ,  $\eta^2 = 0.98$ ). The result of the Bonferroni correction test revealed that there is a significant difference between the pretest, and immediate retention, pretest, and delay retention ( $p < 0.05$ ), whereas there is no significant difference between immediate and delay retention. ANOVA showed a significant main effect for the time in the visual group ( $F_{(2, 38)} = 77.34$ ,  $P = 0.001$ ,  $\eta^2 = 0.79$ ). The result of the Bonferroni correction test revealed that there is a significant difference between pretest and immediate retention, pretest, and delay retention, immediate, and delay retention ( $p < 0.05$ ).

At the pre-test, there was no significant difference between the two groups in the response accuracy. However, there was a significant difference between groups ( $p < 0.05$ ) in the immediate and delay retention (Fig. 3).

### 3.3 Mean fixation number and fixation duration

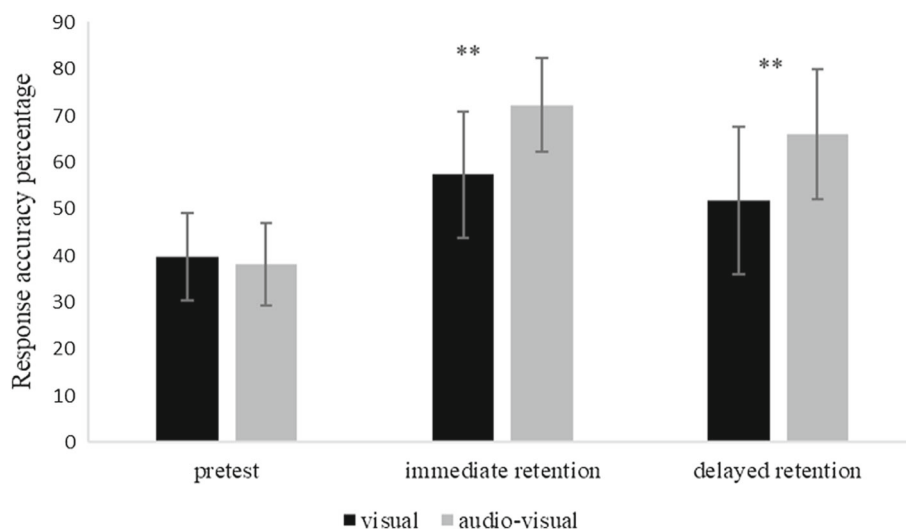
All dependent variables in search rate (mean fixations number and fixation duration) were analyzed using ( $2 \times 3 \times 2$ ) repeated measures ANOVA with response accuracy (correct, incorrect) and time of measurement (pretest, immediate retention, delayed retention) as a within-subject factor and groups (visual, audio-visual) as a between-subject factor.

Analysis of mean fixation durations showed significant main effect for audio-visual group ( $F_{(2, 38)} = 513.82$ ,  $p < 0.001$ , and  $\eta^2 = 0.96$ ) and for mean fixation number ( $F_{(2, 38)} = 235.91$ ,  $p < 0.001$ , and  $\eta^2 = 0.92$ ).

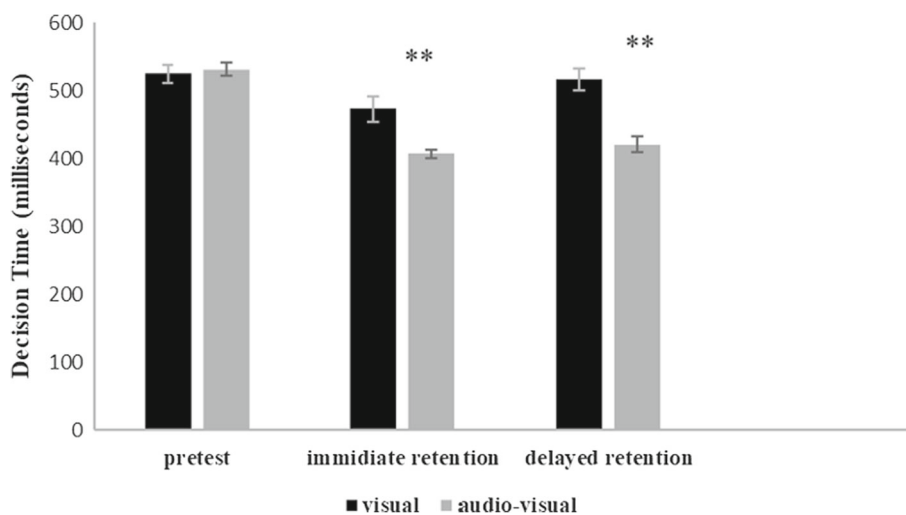
Performance with audio in the adequate response was both associated with a lower number of fixations and longer fixation durations. The audio-visual group made  $460.05 \pm 6.39$  ms of fixation duration with  $21.73 \pm 2.41$  number of fixations in immediate retention and  $457.30 \pm 7.31$  of fixation duration with  $21.83 \pm 2.42$  number of fixations in delayed retention for the correct response.

Percentage Viewing Time. A  $2 \times 5 \times 2$  repeated-measures ANOVA was used to analyze percentage viewing time in which response accuracy (correct, incorrect) and fixation locations (head, trunk, feet, hand, and racket) were the within-subjects factor and groups (visual and audio-visual) the between-subject factors. The analysis indicated significant main effects for groups ( $F_{(2, 38)} = 192.84$ ,  $P < 0.001$ , and  $\eta^2 = 0.89$ ), fixation location ( $F_{(4, 152)} = 408.20$ ,  $P < 0.001$ , and  $\eta^2 = 0.91$ ), and accuracy ( $F_{(2, 38)} = 347.49$ ,  $P = 0.001$ ,  $\eta^2 = 0.95$ ). Analysis also showed groups  $\times$  accuracy ( $F_{(2, 38)} = 87.41$ ,  $P < 0.001$ ,  $\eta^2 = 0.78$ ) and fixation location  $\times$  accuracy  $\times$  groups ( $F_{(4, 152)} = 58.46$ ,  $P < 0.001$ ,

**Fig. 2** Response accuracy (%) between groups (visual; audio-visual) across different times (Pretest; Immediate retention; Delayed retention)



**Fig. 3** Decision time (milliseconds) between groups (visual; audio-visual) across different times (Pretest; Immediate retention; Delayed retention)



$\eta^2 = 0.69$ ) interactions. *T*-tests with Bonferroni correction revealed that when responses of the shot were incorrect, the audio-visual group showed longer fixation on hand ( $t = 1.98$ ;  $P < 0.001$ ) area in comparison to visual group in immediate and delayed retention (see Fig. 4a). Paired sample *t*-tests, with Bonferroni correction, showed that only audio-visual group revealed significant differences between areas on correct/incorrect trials looking longer on the trunk ( $t = 10.03$ ;  $P < 0.001$ ) and hand ( $t = 8.64$ ;  $P < 0.001$ ) areas when the response was correct than incorrect (see Fig. 4b) and rather than the visual group. However, the visual group indicated a significant difference in correct and incorrect responses in racket area at immediate and delayed retention ( $t = 12.63$ ;  $P < 0.001$ ).

### 3.4 Questionnaire answers by the subjects

On average, the visual group was rated as less comfortable than the audio-visual group. Both groups rated their training

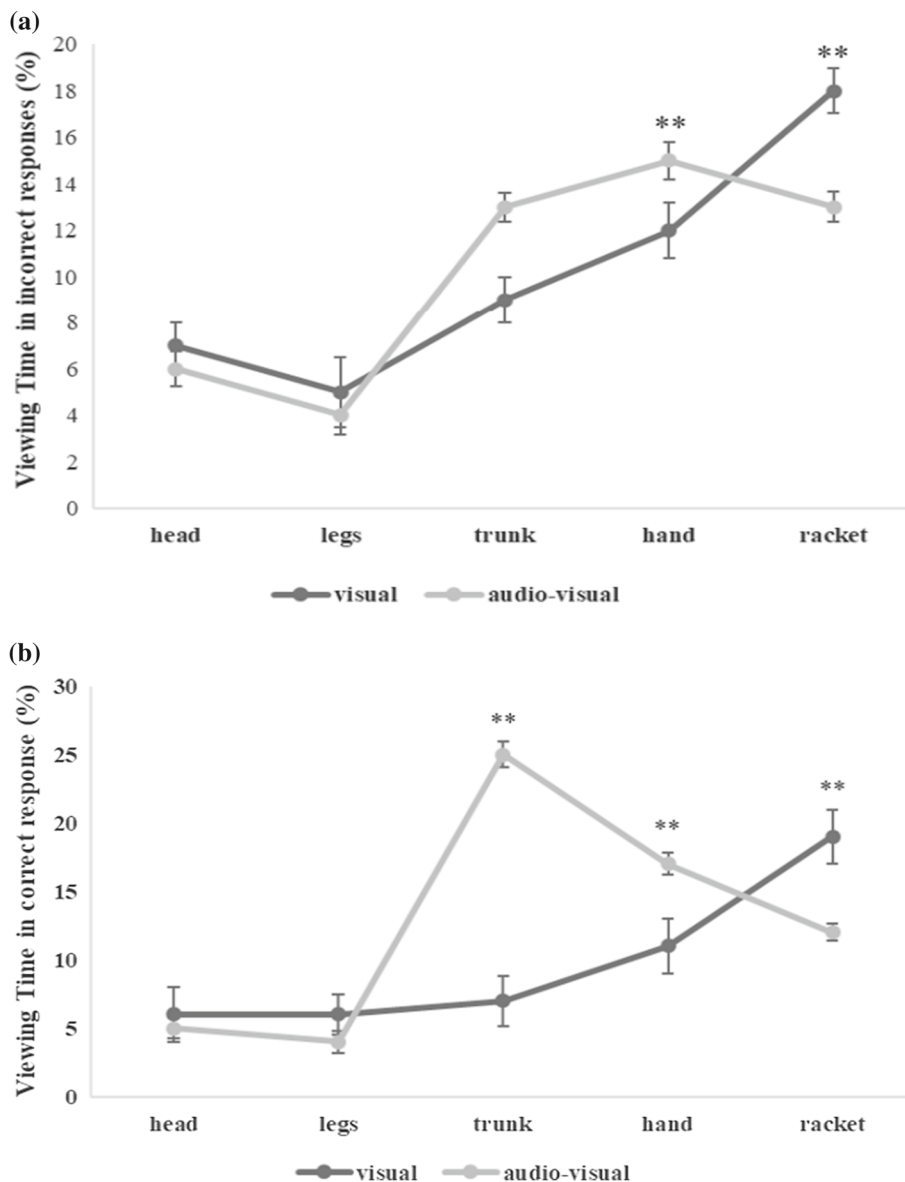
to be similarly useful, and the response of the groups showed that the audio-visual group was more active than the visual group (Fig. 5).

## 4 Discussion

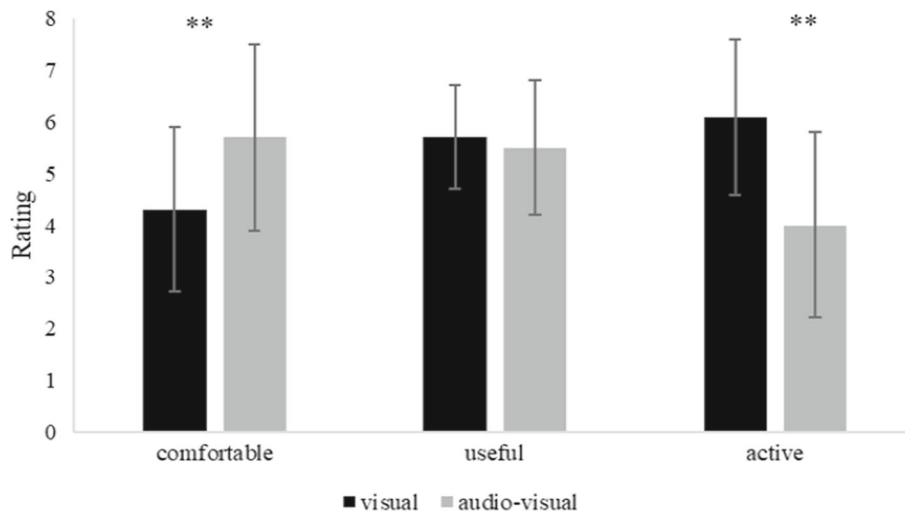
This study aimed to determine whether visual search patterns and anticipation skills can be enhanced through eye movements sonification with instruction relating to the important information cues underlying successful performance. Participants completed a temporal occlusion anticipation test while wearing a mobile eye tracking device with different methods. Results revealed that the audio-visual group had better response accuracy and decision time than the visual group. Moreover, the audio-visual group showed longer fixation durations in key areas.

The findings of the present study extended previous research by showing that novices made the wrong anticipa-

**Fig. 4 a** Percentage viewing time spent by participants on each fixation locations for the incorrect responses between groups. **b** Percentage viewing time spent by participants on each fixation locations for the correct responses between groups. Asterisks indicate significant differences at  $p < 0.05$



**Fig. 5** Questionnaire responses characteristics across the two groups. Asterisks indicate significant differences at  $p < 0.05$



tion when they fixated on the incorrect kinematic locations. Alder et al. (2014) found that expert players fixated on the specific locations that were key factors for anticipation and decision-making in badminton-shot. In their study, novice players also made more incorrect judgments compared to experts when they fixated on the shuttle and other locations, which did not differentiate serve types [4]. The results of gaze behavior in our audio-visual group showed that they learned to pick up subtle differences in postural cues and ignore other irrelevant information. The visual group, that was guided explicitly and directly during training, probably not only they did not know the effectiveness of using the eye movement pattern but also, the directed instruction did not help them during the delayed retention test. Observed changes in the audio-visual group performance implied that this is a meaningful improvement rather than increased test familiarity or habituation. The positive effects of sound coupled with visual instruction support and extend previous studies attempts to enhance perceptual-cognitive skill in sport and other settings [7, 15, 22]. When participants were asked about the type of training preferred, the audio-visual group expressed that training with acoustic signal was attractive and decreased their fatigue [16].

At a broader level, our findings are consistent with a paradigm of sensory processing in which perceptual and cognitive mechanisms are tuned to process multisensory signals. Under such a regime, encoding, storing, and retrieving perceptual information is intended by default to operate in a multisensory environment, and unisensory processing is sub-optimal because it would correspond to an artificial mode of processing that does not utilize the perceptual machinery to its fullest potential [20]. As Buta and Aura (2017) expressed multisensory processes facilitate perception of stimuli and can likewise enhance later areas recognition. Memories for objects and areas originally encountered in a multisensory context can be more robust than those for areas encountered in an exclusively visual or auditory context. In this study, we showed that trial exposures to multisensory contexts are sufficient to improve recognition performance relative to purely unisensory contexts. Multisensory training (audio-visual) probably promotes more effective learning of the information than unisensory training. Although these findings span a large range of processing levels and might be mediated by different mechanisms, it nonetheless seems that the multisensory benefit to learning is an overarching phenomenon.

The findings of this study are supported by the research of Hagman et al. (2006). They examined the anticipatory skills in novice and expert athletes using perceptual training by orienting visual attention using red light. Results of their research showed that badminton novices who trained with red light significantly improved their anticipatory skill between posttest and delayed retention test compared with controls

[7]. In the present study, perceptual training by guiding gaze behavior through sound improved the anticipatory skill of novice participants. One of the reasons for the superiority of the audio-visual group was the use of the guided exploratory learning method to learn key points for anticipation. Participants heard an unpleasant sound when they moved away from the reference points and a pleasant sound when they approached the key points, while they were told before the start of the exercise what situation a pleasant and unpleasant sound they would hear. In this regard, Smeeton et al. (2005) also used the relative effectiveness of a different instruction to improve anticipation skills. Although they stated training facilitated anticipation skill, guided discovery methods are recommended for expediency in learning and resilience under pressure irrespective of the type of instruction used in the experiment. In the present study, guided exploratory (audio-visual) lead to better decision time and higher accuracy [31].

The other result of the present study supported previous findings that suggested fewer number and longer duration of fixations, which indicates skillful performance [4, 13]. In this study, the participant's eye movements convert to sound and transformed into acoustic signals for guiding participants to the key areas of the body, leading to longer fixation durations and fewer fixation numbers. Since each of the two groups improved their gaze behavior patterns, this indicates that participants could learn the exact distinction between different places in the posture from pretest to immediate and delay retention. Then, the audio-visual group can detect the result of different badminton shot and ignore irrelevant information. It can be highlighted that this type of exercise (audio-visual) brought the gaze behavior of the audio-visual group closer to the gaze behavior of expert athletes. Another important consequence of the present study is that augmented sensory could be displayed acoustically to minimize perceptual overload. Using multisensory learning may not only reallocate perceptual and cognitive workload but also reduce distraction, since, unlike visual perception, auditory perception requires neither specific athlete orientation nor a focus of attention [16, 29, 30]. When sonification was withdrawn, participants' performance remained stable during the delayed retention test. In contrast, a second group of participants who had practiced with just visual stimuli showed a decline in performance during the delayed retention test. In this task, probably, sonification preserved the spatio-temporal structure of the task in the perceptual-motor workspace, acting as a guide for pickup main information. Conversely, unisensory training (visual) provides information for the direct control of gaze behavior; its removal means that guidance was no longer possible as the required information was absent [32]. If learning is seen as education of attention [33], the need to sonify task-intrinsic events to avoid the guidance effect is clear [32].



However, the result of the present study was incompatible with Coutrot and et al. (2012). In their study on soundtracks in videos, it was concluded that sound might influence eye position, fixation duration, and saccade amplitude, but that the effect of sound is not necessarily constant over time, whereas in the present study, the acoustic signal guide eye movements during task performance. Probably a factor for the contradictory result was short time duration of our study, whereas the time duration in the Coutrot et al.'s study was too long [24]. Presumably, the audio-visual training group watched longer all interest areas like fixating on trunk and hand when they gave correct responses, indicating that they were trying to foresee player's intention. Longer fixation durations and lower fixation numbers allow the audio-visual training group to spend more time on extracting information from the opponent's kinematic points of posture. Probably, this group, like the gaze behavior in experts could reduce the amount of information to be processed or requires fewer fixations to create a coherent perceptual representation of the display [13].

According to the results of the questionnaire, probably multisensory training can be considered more useful than visual exercise. Since this group stated that they feel comfortable in it and they considered themselves more active. Multisensory training may have increased participants' motivation to use the sound during practice sessions, leading to the increased effort during the training period. As participants' performance decreases during fatigue, it is suggested that the role of eye movement sonification as a motivator for predicting the shot within mental and physical fatigue should be considered. In general, the use of the multisensory learning is likely to be effective in learning complex motor tasks, facilitating the discovery of the new task needs and helping to perceive the exercise structure.

**Acknowledgements** We wish to thank athletes and sport students for taking part in the testing. This research received grant from funding agencies in Shahid Beheshti University.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest regarding the publication of this paper.

## References

- Piras A, Lanzoni IM, Raffi M, Persiani M (2016) The within-task criterion to determine successful and unsuccessful table tennis players. *Int J Sports Sci Coach* 11(4):523–531
- Piras A, Raffi M, Perazzolo M, Lanzoni IM (2019) Microsaccades and interest areas during free-viewing sport task. *J Sports Sci* 37:980–987
- Timmis MA, Piras A, van Paridon KN (2018) Keep your eye on the ball; the impact of an anticipatory fixation during successful and unsuccessful soccer penalty kicks. *Front Psychol* 9:2058
- Alder D, Ford PR, Causer J (2014) The coupling between gaze behavior and opponent kinematics during anticipation of badminton shots. *Human Movement Sci* 37:167–179
- Loffing F, Cañal-Bruland RJCOiP (2017) Anticipation in sport, vol 16, pp 6–11
- Williams AM, Jackson RCJPos, Exercise (2019) Anticipation in sport: Fifty years on, what have we learned and what research still needs to be undertaken? vol 42, pp 16–24
- Hagemann N, Strauss B (2006) Training perceptual skill by orienting visual attention. *J Sport Exercise Psychol* 28(2):143–158
- Abernethy B (1987) The relationship between expertise and visual search strategy in a racquet sport. *Human Movement Sci* 6(4):283–319
- Prigent E, Hansen C, Baurès R, Darracq C (2015) Predicting where a ball will land: from thrower's body language to ball's motion. *Exp Brain Res* 233(2):567–576
- Triplet C, Benguigui N, Le Runigo C (2013) Quantifying the nature of anticipation in professional tennis. *J Sports Sci* 31(8):820–830
- Sáenz-Moncaleano C, Basevitch I (2018) Gaze behaviors during serve returns in tennis: A comparison between intermediate-and high-skill players. *J Sport Exercise Psychol* 40(2):49–59
- Piras A, Raffi M, Lanzoni IM, Persiani M, Squatrito S (2015) Microsaccades and prediction of a motor act outcome in a dynamic sport situation. *Invest Ophthalmol Vis Sci* 56(8):4520–4530
- Piras A, Lobietti R, Squatrito SJJoO (2014) Response time, visual search strategy, and anticipatory skills in volleyball players
- Papin JP, Metges PJ, Amalberti RR (1984) Use of NAC eye mark by radiologists. *Adv Psychol* 22:323–330
- Broadbent DP, Causer J, Williams AM (2015) Perceptual-cognitive skill training and its transfer to expert performance in the field: future research directions. *Europ J Sport Sci* 15(4):322–331
- Losing V, Rottkamp L, Zeunert M, Pfeiffer T, editors (2014) Guiding visual search tasks using gaze-contingent auditory feedback. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*
- Bota A (2017) Multisensory learning in motor activities. *Sci Bull Mircea cel Batran Naval Acad* 20(1):551
- Frid E, Bresin R, Pysander E-LS, Moll JJS, Computing M (2017) An exploratory study on the effect of auditory feedback on gaze behavior in a virtual throwing task with and without haptic feedback, pp 242–249
- Cañal-Bruland R, Muller F, Lach B, Spence C (2018) Auditory contributions to visual anticipation in tennis. *Psychol Sport Exercise* 36:100–103
- Seitz AR, Kim R, Shams LJCB (2006) Sound facilitates visual learning. *Current Biol* 16(14):1422–1427
- Min X, Zhai G, Gao Z, Hu C, Yang X, editors (2014) Sound influences visual attention discriminately in videos. In: *2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*, IEEE
- Klatt S (2020) Visual and auditory information during decision making in sport. *J Sport Exercise Psychol* 42(1):15–25
- Shams L (2008) Benefits of multisensory learning. *Trends Cognit Sci* 12(11):411–417
- Coutrot A, Guyader N, Ionescu G, Caplier A (2012) Influence of soundtrack on eye movements during video exploration
- Cañal-Bruland R, Müller F, Lach B (2018) Auditory contributions to visual anticipation in tennis. *Psychol Sport Exercise* 36:100–103
- Abernethy B, Gill DP, Parks SL, Packer STJP (2001) Expertise and the perception of kinematic and situational probability information. *Perception* 30(2):233–252
- Parseihian G, Aramaki M, Ystad S, Kronland-Martinet R, editors (2017) Exploration of sonification strategies for guidance in a blind

- driving game. In: International symposium on computer music multidisciplinary research, Springer
28. Abernethy B (1987) Expert-novice differences in an applied selective attention task. *J Sport Exercise Psychol* 9(4):326–345
  29. Sigrist R, Rauter G, Marchal-Crespo L, Riener R (2015) Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Exp Brain Res* 233(3):909–925
  30. Sigrist R, Rauter G, Riener R (2013) Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic Bull Rev* 20(1):21–53
  31. Smeeton NJ, Williams AM, Hodges NJ (2005) The relative effectiveness of various instructional approaches in developing anticipation skill. *J Exp Psychol Appl* 11(2):98
  32. Dyer JF, Stapleton P (2017) Mapping sonification for perception and action in motor skill learning. *Front Neurosci* 11:463
  33. Jacobs DM (2007) Direct learning. *Ecol Psychol* 19(4):321–349

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.