Int. J. Sport Psychol., 2023; 54: 248-268
doi: 10.7352/IJSP.2023.54.248

Perceptual training in Beach-Volleyball defence: no beneficial effects of colour-cue interventions

ANDRÉ KLOSTERMANN, *and* ERNST-JOACHIM HOSSNER

Institute of Sport Science, University of Bern, Bern, Switzerland

The colour cue method has been proposed as a perceptual-training intervention in sports. However, the empirical evidence is ambiguous, possibly ascribed to an insufficient match between training and test conditions. Thus, in the test phases of Experiment 1, participants responded either verbally or with a mimicked action to beach-volleyball attacks after having trained with colour-cued gaze paths that were extracted from experts whom them selves had to respond verbally or actively. In Experiment 2, conditions were further matched by making participants mimic actions already over the intervention phase. In contradiction to the expectation that learning-enhancing effects appear for perfectly matched training-testing conditions at least, no differences were revealed, neither between the colour-cue in terventions nor between colour-cueing and the control condition of just watching the same videos. However, gaze was reliably affected by response modes, meaning that gaze behaviour substantially changes if either verbal or active responses are required.

KEY WORDS: Perceptual-cognitive expertise, Decision making, Gaze learning, visual attention, Perception.

Perceptual-cognitive skills facilitate the processing and organization of environmental information when executing appropriate actions (Marteniuk, 1976). Generally, it is found that skilled athletes show superior perceptual-cognitive skills as evidenced by better anticipation and decision-making performance as well as more efficient and effective gaze behaviour which results in superior task completion (Williams et al., 2011; for an overview, e.g., Brams et al., 2019; Mann et al., 2007). This means that in complex situations, like an attack in beach volleyball, which require multiple potential actions from the defender, the skilled defender can better anticipate the most probable type of attack and make the right decision to defend the attack successfully. Hence, from a practical point of view, it is consequent to strive

---

Correspondence to: André Klostermann, Bremgartenstrasse 145, CH-3012 (E-mail: andre.klostermann@unibe.ch)

to enhance those skills through respective training methods (for an overview, Larkin et al., 2015). In this regard, video-based perceptual training programs have been suggested to acquire and refine these skills (for a classification Hadlow et al., 2018). Although existing research implies a general effectivity of these programs (e.g., Larkin, et al., 2015) as well as – to a certain degree – transfer into practice (e.g., Williams et al., 2011), the optimal form of this treatment is still under debate (Abernethy et al., 2012; Broadbent et al., 2015). For example, Fleddermann et al. (2019) investigated the effects of an eight-week multiple-object tracking intervention to (rather implicitly) train visual perception by having the learners monitor multiple (virtual) objects (Neurotracker trainig). Results, however, showed improvements only for task-specific (i.e., the multiple-object tracking task) and near-transfer tasks (cognitive tests) but not for decision-making in a volleyball-specific blocking test (for a recent review on the Neurotracker training tool, Vater et al., 2021).

Another prominent candidate for a more explicit gaze-related perceptual training to enhance decision-making skills is the so-called colour-cueing method that draws on the classical spatial-cueing paradigm introduced by Posner (1980). By implementing salient visual patches into training videos, it is expected that the learners' point of gaze shifts to the highlighted information sources, thereby reducing the amount of information to be processed due to optimised attentional control. The target gaze path is commonly derived from experts' gaze behaviour (e.g., Hagemann et al., 2006). Thus, it is assumed that adapting the experts' gaze path will improve decision-making performance (Mann et al., 2007; for an overview, Gegenfurtner et al., 2011). Consequently, when guiding beginners' visual attention to the most information-rich areas, processing this information shall be facilitated and result in improved decision-making.

However, the benefit of this perceptual gaze training for decision-making delivered somewhat inconsistent results. On the one hand, enhancing effects of colour-cueing interventions have been reported for ball-flight anticipation in badminton by Hagemann, et al. (2006) as well as for the anticipation of the penalty kicking direction in football by Savelsbergh et al. (2010) and by Ryu et al. (2013). On the other hand, in a direct comparison of different perceptual-learning protocols for improving handball goalkeepers' skill to predict the shot direction, Abernethy et al. (2012, p. 143) summarise as follows: "The explicit learning, verbal cueing, and implicit learning conditions provided the greatest sustained improvements in performance whereas the group given colour cueing performed no better than the control groups.". Moreover, Cañal-Bruland (2009) showed that – with nearly the same visual stimuli – the cueing paradigm facilitates performance in signal-detection tasks but not in decision-making tasks.

Given this inconsistent situation, Klostermann et al. (2015) strived to get a closer grip on the colour-cueing approach by introducing eye-tracking measurements, thereby experimentally disentangling the effects of colour cueing on gaze behaviour on the one and on anticipation performance on the other hand. First of all, in a preparatory study on cue evaluation, they found earlier fixation onsets in a volleyball padding-direction anticipation task, accompanied by higher response accuracies, in particular for flickering and comparably large colour cues such that these kinds of cues were used in the main study on the anticipation of attack directions in beach-volleyball defence. In this study, a video-based training without cueing was contrasted with two cueing interventions, namely a functional protocol with a gaze path derived from a top-level player (with saccades from the ball to the attacker and from the attacker to the anticipated location of the ball-hand contact) and a dysfunctional protocol in which the ball was continuously cued. As expected, a considerable difference in gaze behaviour between the functional and the dysfunctional group was revealed over the acquisition phase and post- and retention tests, showing that the intervention effectively altered gaze behaviour. However, this difference was not reflected in superior anticipation scores for the functional group since no significantly different learning rates could be found between the colour-cue groups. The control group without cueing improved most from pre- to post- and retention test. Hence, the results reported by Klostermann et al. (2015) corroborate the results of the Abernethy et al. (2012) study and give further reason to question the effectiveness of colour-cueing methods for decision-making training in sports (see also Cañal-Bruland, 2009).

However, before finally accepting this conclusion, some precaution seems advisable. In the Klostermann et al. (2015) study, the lack of empirical support for colour-cueing interventions could be ascribed to an insufficient match of training and test conditions. This issue regards the fact that participants had to learn gaze paths derived from test conditions in which experts initiated an actual movement in the direction of the attack. However, in the pre-, post- and retention tests of the colour-cue study, a temporal-occlusion paradigm was applied, meaning that the participants verbally specified their prediction immediately after occlusion rather than initiating a motor response. This difference is crucial as experts when initiating a motor response, typically show the above sketched gaze-path pattern with a final saccade from the attacker to the anticipated location of the ball-hand contact. In contrast, they regularly focus on the attacker under verbal response conditions (for this context, the highly relevant difference between vision-for-action and vision-for-perception conditions, e.g., Dicks et al., 2010).

To resolve this mismatch, the test conditions should be adapted to the respective gaze-intervention protocol (i.e., verbal test conditions after a verbal gaze-path protocol and action test conditions after an action gaze-path protocol, respectively). Beyond, to increase the chance for revealing learning-enhancing effects of gaze-path colour cueing even further, the interventions should be adapted to the test conditions regarding the tested gaze paths and the – either active or verbal – responses that are required from the participants. In a stepwise approach, the first of these steps will be taken in Experiment 1 and the second in Experiment 2.

Experiment 1: Matched Test Conditions

In Experiment 1, we drew on the experimental design by Klostermann et al. (2015) on anticipatory-skill training in beach-volleyball defence. As in the original study, participants received a video-training intervention to learn experts' gaze behaviour. The experimental group was trained with colour cues derived from top-level players' gaze behaviour when performing under action conditions (denoted as S-action with S for "stimulus"). The control group was trained with the same videos but without colour cues (denoted as S-no). Unlike the original study, these two groups were not contrasted with a dysfunctional colour-cue group (i.e., highlighting the ball flight) but with a group trained with colour cues derived from top-level players when verbally responding to occluded scenes (denoted as S-verbal). Furthermore, in comparison to Klostermann et al. (2015), the occlusion test with a verbal specification of the upcoming type of attack (denoted as T-verbal with T for "test") was complemented by a test in which the appropriate defence movement should be initiated (denoted as T-action). Hence, one ended up with three groups labelled after their treatment as S-action, S-verbal and S-no, respectively, whose pre-, post- and retention test performance was measured in two tests, namely T-action and T-verbal. It was predicted that the colour-cue groups would outperform the S-no group regarding their improvements from pre- to posttest and pre- to retention test, particularly under matched test conditions (i.e., S-action in T-action and S-verbal in T-verbal).

Method

PARTICIPANTS

23 male (age: 22.3 ± 3.3 years) and 19 female (age: 21.3 ± 0.8 years) sport science students participated in the study, receiving course credits in return. The participants were as-

signed to one of three intervention groups to secure comparable pretest scores based on their pretest decision-making performance and gaze behaviour. All participants had self-reported normal or corrected-to-normal vision and were unaware of the research question. The experiment was undertaken in accordance with the 1964 Declaration of Helsinki.

Video scenes. The raw video scenes were taken from a comprehensive video database on expert beach-volleyball attacks from the perspective of a cross-court defence player with a teammate blocking longline at the right side of the net. The two opposing players performed standardised attacking sequences, composed of reception, set and an attack that was executed as a hard cross-court smash towards the defender's position, a line shot over the block to the (from the defender's perspective) court's right-back corner, or a cut shot to the (from the defender's perspective) court's left-front corner. Thus, the real-world task would require the defender to keep his or her position against cross-court smashes or run either right-backwards or left-forwards to reach line shots and cut shots, respectively.

The video scenes showed two male and two female, right-handed and internationally experienced attackers. From 144 raw scenes, for each of the four players, four scenes per type of attack were selected both for the test and intervention sessions, which either were presented untreated to the S-no control group or had to be further processed for the colour-cue interventions. For the latter scenes, the beforehand acquired experts' gaze behaviour, either in the action or verbal condition was highlighted by coloured patches using a self-written Matlab 2014b routine (cf. Klostermann et al., 2015). In more detail, as illustrated in Figure 1, all video scenes were first subdivided into two consecutive phases, phase P1 for the run-up from the set until the initiation of the attacker's jump (defined by the beginning of the upswing of the arms) and phase P2 for the attack from the end of the previous phase until the attacker's ball-hand contact. Over P1, the colour cues always followed the attacker's upper body. Still, they differed for the two colour-cue groups over P2, thereby (in an ideal manner) reflecting experts' gaze behaviour when watching video clips under either T-action or T-verbal test conditions. This means that the colour-cue patches jumped to the predicted location of the attacker's ball-hand contact in the S-action intervention, whilst a jump to the attacker's right arm was initiated in the S-verbal intervention.

After editing the video scenes, three sets of 24 scenes each were available for the S-action, S-verbal and S-no intervention. These (non-occluded) scenes were rendered into 12 blocks of 12 video scenes, meaning each scene appeared six times (two times each in blocks 1-4, 5-8, and 9-12, respectively). Each block comprised each type of attack four times, arranged in a random order (MAGIX Video Pro X3). As the tests following the intervention phase had to be performed without gaze-path cues, the frequency of cued scenes was gradually reduced over the intervention phase for the colour-cue groups (blocks 1-4: 100% cued scenes, blocks 5-8: 67% cued scenes, blocks 9-12: 33% cued scenes) by replacing the cued scenes by the respective scenes without colour cues (cf. Klostermann et al., 2015).

For the T-action test, the video scenes (without colour cues) were rendered into two blocks of 12 trials each in a quasi-randomised order (each block containing each type of attack four times) such that each scene was presented exactly once in pre-, post- and retention test. For the T-verbal test, these scenes were further processed by replacing the frames from 200 ms before the attacker's ball-hand contact until the end of the video by blue frames, thereby occluding the actual attacking movement. Finally, audio triggers were added at the beginning and end of each scene's crucial phases P1 and P2, to relate participants' behaviour to the video footage.

Apparatus. Participants' gaze behaviour was recorded with a VICON-integrated mobile eye-tracking system (EyeSeeCam, 220Hz) that was connected to a MacBook Pro via a 20 m fibre-optic Fire Wire link (GOF-Repeater 800, Unibrain). To allow the participants to move freely in the laboratory they wore a bum bag which stored the repeater and the power supply. The integrated EyeSeeCam system uses the infrared corneal reflection method to assess the vertical and horizontal rotations of both eyes. Three-dimensional translations and rotations of the participant's head were derived by the VICON system from the positions of three retro-reflective markers attached to the EyeSeeCam, allowing for the calculation of a three-dimensional gaze vector that is updated every 5ms (Kredel et al., 2011). The accuracy of the eye-tracking system amounts to 0.5° of the visual angle, with a resolution of 0.01° root-mean squared error, within 25° of the participant's field of view. The EyeSeeCam was (re)calibrated at the beginning and in the mid of each test session. The gaze was only recorded in pre-, post-, and retention tests but not for acquisition trials. Data processing was executed with Mathworks MATLAB 2013a, and IBM SPSS Statistics 22 was used for statistical analyses.

For obtaining measures characterising decision making in the T-action test, the type of indicated defence movement was put down in writing by an experimenter, whilst the moment of movement initiation was detected by two force plates (AMTI, 1000Hz) as well as by the motion-capture system (VICON, 220Hz). In the T-verbal test, participants verbally responded to the upcoming type of attack, and the anticipations were put down in writing by an experimenter.

PROCEDURE

The study was conducted in the institute's sensorimotor laboratory, and participants attended three individual sessions. After having read the instructions, they were fitted with the EyeSeeCam followed by a calibration that required to consecutively fixate five dots displayed in a regular grid with a distance of 8.5° of visual angle between the dots (Kredel et al., 2011). Subsequently, either the test or the intervention started. In all sessions, the video scenes were displayed at a life-sized screen (height: 2.4 m, width: 3.6m), and the participants were positioned on the force plates according to the defender's real cross-court position at a distance of 4.0m to the screen. The force plates were (re)calibrated at the beginning of each block in the T-action tests.

In the pretest session 1 and retention-test sessions 3, following a warm-up block of 6 trials (2 types of attacks x 3 repetitions), the T-action and T-verbal tests were conducted with two blocks of 12 test trials each (4 attackers x 3 types of attack x 2 trials) in counter-balanced order. In the T-action test, participants were advised to "defend" the upcoming attack as in a real-game situation by either indicating a whole-body movement in the respective direction (line shot: right-back; cut shot: front-left) or by holding the position (cross-court smash). After each trial, participants returned to the start position on the force plates. The explicit instruction was to decide as early and correct as possible, leaving the respective weighting of these interdependent factors to the participants. In the T-verbal test, participants watched the temporally occluded videos in an upright-standing cross-court position. They were asked to call out the type of attack they would have to defend immediately after the occlusion.

In the beginning of the intervention/posttest session 2, the group-specific intervention consisted of 12 blocks with 12 trials each. Participants were instructed to learn the gaze path depicted by the coloured patch (S-action and S-verbal groups) or to get an idea about the presented type of attacks by watching the presented scenes (S-no group). Beyond, the colour-cue groups were informed about the rationale of the respective cued gaze path. However, neither active nor verbal responses were required from the participants, and thus no feedback was provided over the intervention. The second session was completed by the posttest, conducted after a short break immediately after the intervention phase and structured as described above for the pre- and retention test.

Pre-, post-, and retention tests lasted about 45 minutes each, and the intervention phase about 120 min. The intervention/posttest session 2 occurred about two weeks after the pretest session 1 (14.4 ± 4.8 days) and the retention-test session 3 after about one further week (6.9 ± 0.3 days). At the end of the final session, the participants were thanked and debriefed about the study's objectives.

MEASURES

Gaze behaviour. Due to technical difficulties in data collection, gaze data from two trials (= 0.03%) had to be excluded from further analysis. For the remaining raw analogue data, after down-sampling from 1000Hz to 200Hz and smoothing the horizontal and vertical eye rotations with an 11-point, 3rd order Savitzky-Golay filter, intersection points between the three-dimensional gaze vector and the screen were calculated to determine the two-dimensional gaze path in the screen frame of reference for each trial. Likewise, the digitised two-dimensional pixel coordinates of the coloured patches (25Hz) were converted into the screen-reference frame and up-sampled to 200Hz by linear interpolation. As the colour-cue coordinates represent the target gaze path, distances could be derived for the test sessions between the participant's actual gaze path and the gaze path cued in the S-action and the S-verbal interventions, respectively. These distances were considered for further analyses for phase P2 only because the cued gaze paths did not differ between both cueing conditions over P1. No colour cues were displayed before P1 and after P2 (cf. Figure 1).

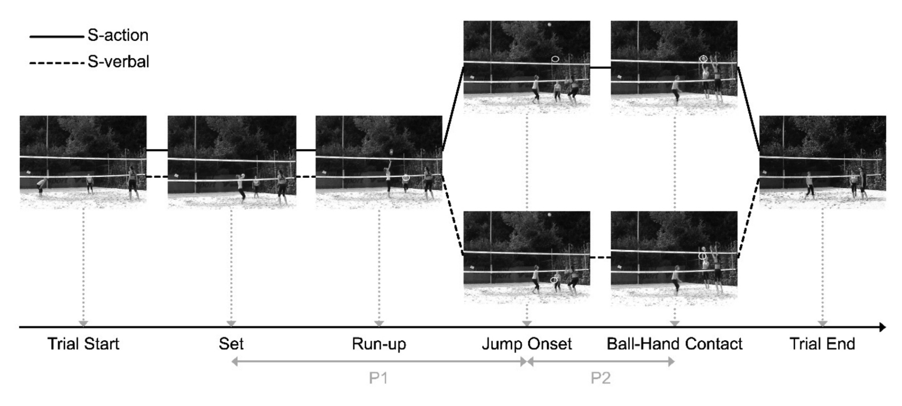

Fig. 1. - Gaze-path learning videos for the S-action and S-verbal intervention, respectively. Whilst over P1 the attacker's upper body was highlighted in both groups, either the predicted ball location at ball-hand contact (S-action) or the attacker's right arm (S-verbal) was high-lighted over P2. Original tapes were coloured with red patches.

In more detail, as introduced by Klostermann et al. (2015), a "gaze-path index" (GPI) was calculated to summarise participants' gaze behaviours in one single dependent variable. To this end, for each frame, the current gaze location was orthogonally projected onto the straight line connecting the momentary coordinates of the S-action and the S-verbal gaze paths. The projection was related to the centre of the line segment (as zero-point) such that the resulting distance between the participant's actual gaze path to the two target gaze paths can be expressed directionally with positive values corresponding to a gaze location closer to the S-action gaze path and negative values corresponding to a gaze location closer to the S-verbal gaze path. The average value over P2 was taken as the dependent gaze measure for the trial. The mean of these values over all valid trials per test and participant was finally converted from screen-related units (mm) into eye-related units of visual angle (°). Consequently, based on the measured average distances between the actual and the two cued gaze paths over P2, GPI values of +2.8° and -2.8° denote that the gaze was exactly on the S-action and the S-verbal gaze path, respectively. In contrast, a value of 0° would correspond to an actual gaze path located exactly in the middle between both target paths.

Decision making. For the calculation of a dependent variable for decision accuracy, participants' responses which had been recorded in writing for the mimicked (T-action) and for the verbally specified response (T-verbal), respectively, were transferred into percentage scores per test and participant (percentage correct out of 24 responses).

A response-initiation measure was also derived from the movement kinematics captured by the VICON system and the force plates for the T-action test. To this end, first, the raw force data were down-sampled to 200Hz and filtered with a 4th order low-pass Butterworth filter (cut-off frequency: 5Hz) (Prieto, Myklebust, Hoffmann, Lovett, & Myklebust, 1996), whereas the raw VICON data for the eye-tracker markers were filtered with a 41 point, 3rd order Savitzky-Golay filter. Second, to calculate the moment of leaving the force plates, the participants' individual starting positions were determined and constantly updated whenever the two-dimensional velocity of the markers fell below a 0.25m/s threshold for at least 500 ms but held fixed as soon as the headmarkers' velocity exceeded 0.25 m/s again. As soon as the current headmarkers' two-dimensional position permanently exceeded a distance of 0.4m in relation to the continuously updated starting position in the transverse plane, this time was taken as the moment the participant left the force plates. Third, starting from this moment, a backwards-in-time search was conducted to detect the first value of the differentiated two-dimensional force vector (in an orthogonal and horizontal direction to the screen) that fell (after the first local maximum) below 1N/s. After relating this time to the moment of ball-hand contact of the attacker, this value was taken as the movement-initiation measure for the trial at hand. Fourth, an outlier elimination was conducted for the resulting 24 raw movement-initiation times per test and participant, excluding all values larger or smaller than *M* ± 2 *SD* (2.9% of all trials). The final response-initiation score was calculated as the mean of the remaining values. Consequently, the finally obtained (generally negative) value for the dependent variable of "relative response initiation time" denotes the interval (in ms) between the point in time when the participant initiates a motor response and the point in time when the attacker hits the ball.

Statistical analyses. After the calculations described above had been conducted, six dependent measures were available for each participant for pre-, post- and retention test: GPI, response accuracy and response initiation for the T-action test, and GPI and response accuracy for the T-verbal test. GPI and response accuracy were subjected to a 3 (group: S-action vs. S-verbal vs. S-no) x 3 (time of measurement: pretest vs. posttest vs. retention test) x 2 (test: T-action vs. T-verbal) mixed-factorial ANOVAs with repeated measures on the last two factors. The response-initiation variable was analysed with a 3 (group: S-action vs. S-verbal vs. S-no) x 3 (time of measurement: pretest vs. posttest vs. retention test) mixed-factorial ANOVAs with repeated measures on the last factor. Significant main and interaction effects were further analysed with planned t-tests. The significance level was a priori set to $\alpha = .05$. A posteriori effect sizes were computed as partial eta squares (η_p^2) and Cohen's *d*. An optimal sample size of $N = 42$ had been determined beforehand based on the expectation of medium effect sizes $\eta_p^2 = .11$ (cf. Klostermann et al., 2015) for an α-level of .05 and a power of .95.

Results

Gaze behaviour. Analyses of the GPI (see Figure 2) revealed a significant main effect for test, $F(2, 78) = 33.89$, $p < .01$, $\eta_p^2 = .47$, with lower GPI scores in the T-verbal (M = -0.47° of visual angle, SE = 0.17° of visual angle) than in the T-action test (M = 0.81° of visual angle, SE = 0.29° of visual angle). Whereas, in the T-verbal condition, participants anchored their gaze on average just below the middle of both target gaze-paths, gaze was clearly shifted towards the S-action path in the T-action condition. All further comparisons, especially the group x time of measurement x test interaction, $F(4,78) = 0.46$, $p > .05$, $\eta_p^2 = .02$, missed the pre-determined level of significance (all ps > .09, all $\eta_p^2 < .07$).

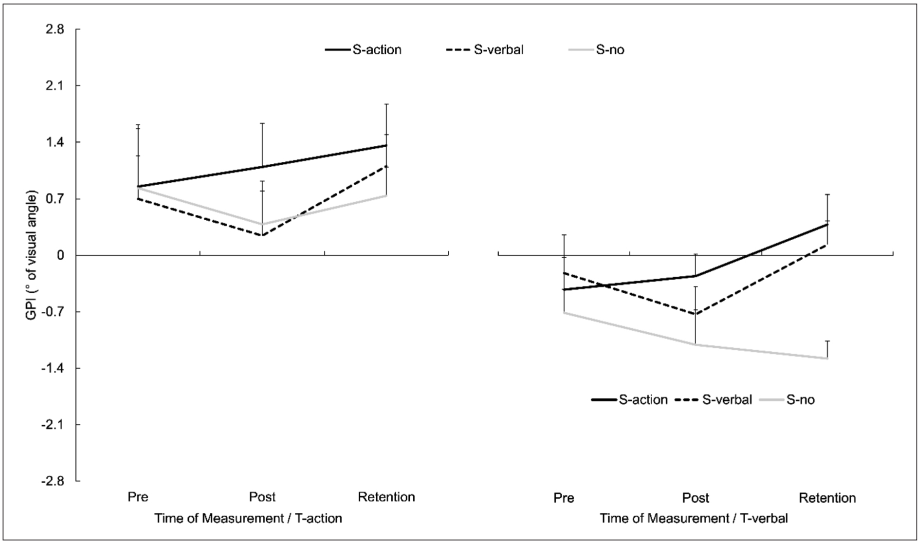

Fig. 2. - Experiment 1: GPI (° of visual angle) as a function of intervention group (S-action vs. S-verbal vs S-no), time of measurement (pretest vs. posttest vs. retention test), and test condition (T-action vs. T-verbal). The value of 2.8° denotes the cued S-action target-gaze path and the value of -2.8 ° the cued S-verbal target-gaze path, respectively.

Decision making. The ANOVA for response accuracy (see Figure 3) revealed a significant main effect for test, $F(2, 78) = 5.38$, $p < .05$, $\eta_p^2 = .12$, which, however, was overlaid by a significant time of measurement x test interaction, $F(2, 78) = 8.44$, $p < .01$, $\eta_p^2 = .18$. Over learning, participants increased response accuracy in the T-verbal test only, $F(2, 82) = 13.55$, $p < .01$, $\eta_p^2 = .25$, with better performance in post- and retention tests than in the pretest (all *p*s < .01, all *d*s > 0.68). In contrast, no learning was found for response accuracy in the T-action test, $F(2, 82) = 0.47$, $p = .63$, $\eta_p^2 = .01$. All further comparison, in particular the predicted three-way interaction ($p = .89$, $\eta_p^2 = .01$), were not significant (all *p*s > .08, all $\eta_p^2 = .08$).

The response-initiation analyses (see Table I, Experiment 1) revealed a significant main effect for measurement only, $F(2, 78) = 26.82$, $p < .01$, $\eta_p^2 = .40$. In both post- and retention test, participants initiated their responses earlier than in the pretest (all ps < .01, all *d*s > 0.83). No further significant main and interaction effects were found (all ps > .11, all $\eta_p^2 < .02$).

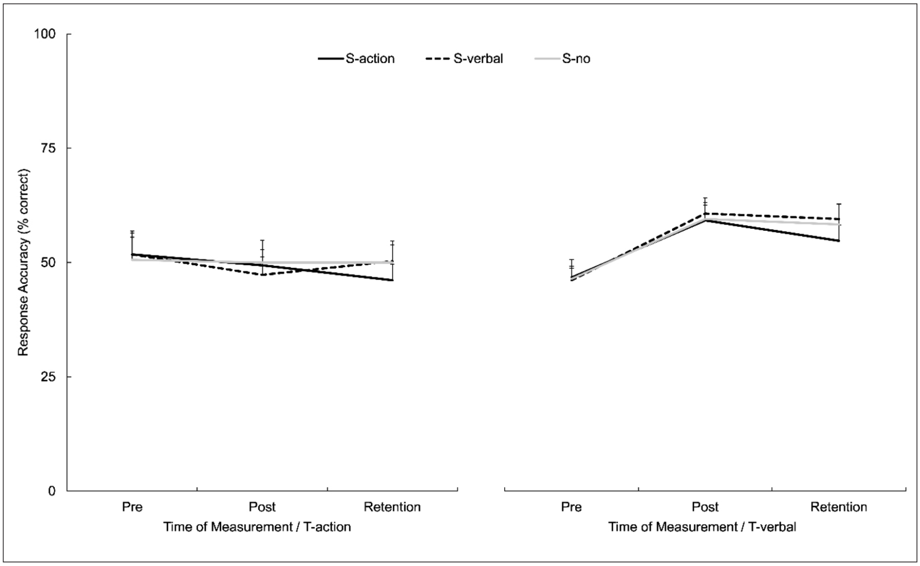

Fig. 3. - Experiment 1: Response accuracy (% correct) as a function of intervention group (S-action vs. S-verbal vs. S-no), time of measurement (pretest vs. posttest vs. retention test), and test condition (T-action vs. T-verbal).

TABLE I
Relative Movement-Initiation Time (Ms) As A Function Of Intervention Groups (S-Action Vs. S-Verbal Vs. S-No) And Time Of Measurement (Pretest Vs. Posttest Vs. Retention test). The value of 0 ms denotes the moment when the attacker hits the ball.

	Group	Movement-Initiation Time (ms)
Pretest	S-action	-102.5 (± 196.7)
	S-verbal	-121.5 (± 268.6)
	S-no	-128.3 (± 219.9)
Posttest	S-action	-361.8 (± 134.6)
	S-verbal	-299.0 (± 216.5)
	S-no	-305.4 (± 222.8)
Retention	S-action	-383.4 (± 136.6)
	S-verbal	-319.9 (± 225.8)
	S-no	-352.6 (± 171.8)

Discussion

In Experiment 1, we tested the effect of colour cueing of two different gaze paths (S-action vs. S-verbal) on gaze behaviour and decision making in two different tests (T-action vs. T-verbal). Regarding decision making, participants improved from pre- to post- and retention test, in the T-verbal test with a higher response accuracy and in the T-action test with an earlier response initiation. However, differences in decision making could neither be revealed between the two colour-cue groups (S-action, S-verbal) nor between colour-cueing and the control condition (S-no). Also gaze behaviour was not differentially affected by the three interventions. Instead, before as well as after the intervention phase, a general effect of the test condition on participants' gaze behaviour was found with participants' gaze paths closer to the experts' gaze path if the respective test conditions (T-action vs. T-verbal) matched the conditions under which the experts' gaze paths had been acquired (S-action vs. S-verbal).

Hence, a strong performance effect of test conditions on gaze behaviour could be revealed, but no enhanced learning due to colour cueing. As this is true for both test conditions, the absence of colour-cueing effects in the Klostermann et al. (2015) study cannot be ascribed to a mismatch between the action-related gaze-path intervention and the verbal test condition.

Experiment 2: Matched Training Conditions

In Experiment 1, the null finding by Klostermann et al. (2015) was not changed by the introduction of gaze-path interventions that matched the test conditions. Thus, as already discussed in the introduction, the question remains whether the same finding can be expected if not only the gaze-path protocol over the intervention was adapted to the test conditions (i.e., S-action, S-verbal and S-no) but also the execution of the respective response over acquisition. Consequently, it was asked in Experiment 2 whether colour-cue interventions would enhance decision making in beach-volleyball defence when the intervention required either active (denoted as R-action) or verbal (denoted as R-verbal) responses.

On this basis, one ends up with a 2x2 design with two group factors: cue (colour cue vs. no colour cue) and response (action vs. verbal). In more detail, in the colour-cue groups, participants either had to mimic defence actions when watching videos in which an experts' gaze path under T-action test conditions was cued (S-action/R-action with R for "response") or they had to verbally specify the predicted type of attack in temporally occluded scenes in which an experts' gaze path under T-verbal test conditions was cued (S-verbal/R-verbal). These two groups were compared with two control groups whose participants practiced without colour cues either with non-occluded videos and mimicking defence actions (S-no/R-action) or with occluded videos and verbally specifying the predicted type of attack (S-no/R-verbal).

Please note that by these specifications, the S-verbal/R-verbal and the S-no/R-verbal group correspond to the S-verbal and the S-no group of Experiment 1, respectively, since no active responses had been required in Experiment 1 as well. However, beyond replicating the results obtained so far (regarding the S-action vs. S-no and the S-verbal vs. S-no comparisons), this design provides a further increased match between intervention and test conditions. Consequently, if colour cueing enhances learning, it should be expected at least that the S-action/R-action in comparison to the S-no/R-action control treatment results in a more pronounced improvement in the T-action test and that the S-verbal/R-verbal in comparison to the S-no/R-verbal control treatment results in a more pronounced improvement in the T-verbal test. Beyond, to increase the chances of finding learning-enhancing effects of colour cueing even further, the duration of the intervention was doubled from one session to two sessions of two hours each. If even these two measures would not at least result in a trend in favour of colour cueing, the overall conclusion seems warranted that the benefit of colour cueing in the context of decision-making training is rather limited.

Method

Participants. 27 male (age: 21.1 ± 1.5 years) and 25 female (age: 20.1 ± 1.8 years) sport science students who did not participate in Experiment 1 were recruited and received course credits in return. The participants were equally assigned to one of the four intervention groups based on their pretest performance. All participants had self-reported normal or corrected-to-normal vision and were unaware of the research question. The experiment was undertaken in accordance with the 1964 Declaration of Helsinki.

Video scenes and apparatus. For the pre-, post- and retention test for all groups and the intervention videos for the action-response groups (S-action/R-action, S-no/R-action), the same video scenes could be taken as in Experiment 1. However, the intervention videos for the two verbal-response groups (S-verbal/R-verbal, S-no/R-verbal) needed to be compiled. To this end, the scenes of the two batteries (either with or without cued gaze path) were occluded three frames (i.e., 120 ms) before the attacker's ball-hand contact. In addition, to even the amount of feedback augmentation in all four intervention groups, the occlusion was removed after 4s, and the video continued showing the actual execution of the attack. The same apparatus was used in test and intervention phases as in Experiment 1.

PROCEDURE

The procedure was the same as in Experiment 1 except for the introduction of verbally given responses in the treatment phase of the groups S-verbal/R-verbal and S-no/R-verbal and the adding of a second intervention session, in which the same videos as in the first intervention session were presented, resulting in a total of 288 practice trials.

The time intervals between the sessions amounted to about two weeks (from pretest to first intervention; 16.2 ± 8.9 days), one week (from first to second intervention/posttest; 7.3 ± 1.1 days), and one further week (from second intervention/posttest to retention test; 6.8 ± 0.8 days).

MEASURES

The same dependent measures were calculated as in the first study for the T-action and T-verbal tests. In Experiment 2, the outlier-elimination procedure for the response-initiation variable resulted in 3.1% of all trials being excluded from further analyses. GPI and response accuracy were subjected to 2 (cue:

S-action/S-verbal vs. S-no) x 2 (response: R-action vs. R-verbal) x 2 (test: T-action vs. T-verbal) x 3 (time of measurement: pretest vs. posttest vs. retention test) mixed-factorial ANOVAs with repeated measures on the last two factors. The response-initiation variable was analysed with a 2 (cue: S-action/S-verbal vs. S-no) x 2 (response: R-action vs. R-verbal) x 3 (time of measurement: pretest vs. posttest vs. retention test) mixed-factorial ANOVA with repeated measures on the last factor. Significant main and interaction effects were further analysed with planned t-tests. A posteriori effect sizes were computed as partial eta squares (η_p^2) and Cohen's *d*. Optimal sample-size calculations were based on the same approach as in Experiment 1. However, due to technical problems with the eye-tracker, the data of 4 participants could not be analysed such that we finally ended up with data sets of 4x13 (rather than 4x14) participants.

Results

Gaze behaviour. Analyses of the GPI (see Figure 4) revealed a significant main effect for test, $F(1, 48) = 18.71$, $p < .01$, $\eta_p^2 = .28$. Participants showed

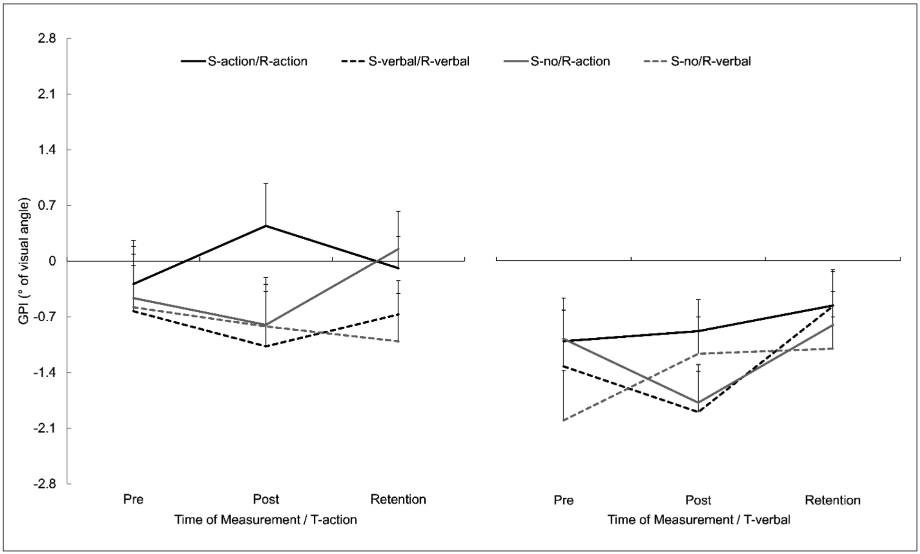

Fig. 4. - Experiment 2: GPI (° of visual angle) as a function of intervention group (S-action/R-action vs. S-verbal/R-verbal vs. S-no/R-action vs. S-no/R-verbal), time of measurement (pretest vs. posttest vs. retention test), and test condition (T-action vs. T-verbal). The value of 2.8° denotes the cued S-action target-gaze path and the value of -2.8° the cued S-verbal target-gaze path, respectively.

different gaze behaviours by anchoring their gaze mainly in-between the S-action and the S-verbal target gaze paths in the T-action test (M = -0.49° of visual angle, SE = 0.23° of visual angle), but closer to the S-verbal target gaze path in the T-verbal test (M = -1.17° of visual angle, SE = 0.19° of visual angle). In addition, a significant three-way interaction cue x response x time of measurement, $F(2, 96) = 3.12$, $p < .05$, $\eta_p^2 = .06$, was found, which, however, did not result in significant follow-up comparisons. No other significant main or interaction effects were found (all ps > .14, all $\eta_p^2 < .04$).

Decision making. The ANOVA on response accuracy (see Figure 5) revealed significant main effects for measurement, $F(2, 96) = 40.62$, $p < .01$, $\eta_p^2 = .46$, and test, $F(1, 48) = 48.57$, $p < .01$, $\eta_p^2 = .50$. On the one hand, participants showed superior performance in the posttest, $t(51) = 7.74$, $p < .01$, $d = 1.21$, and the retention test, $t(51) = 7.74$, $p < .01$, $d = 1.13$, when compared to the pretest, but no significant differences between post- and retention test, $t(51) = 0.85$, $p = .39$, $d = 0.12$. On the other hand, participants responded more accurately in the T-verbal test (M = 57.7%, SE = 1.2%) than in the T-action test (M = 46.7%, SE = 1.2%). No other significant main and interaction effects were found (all ps > 14, $\eta_p^2 < .05$).

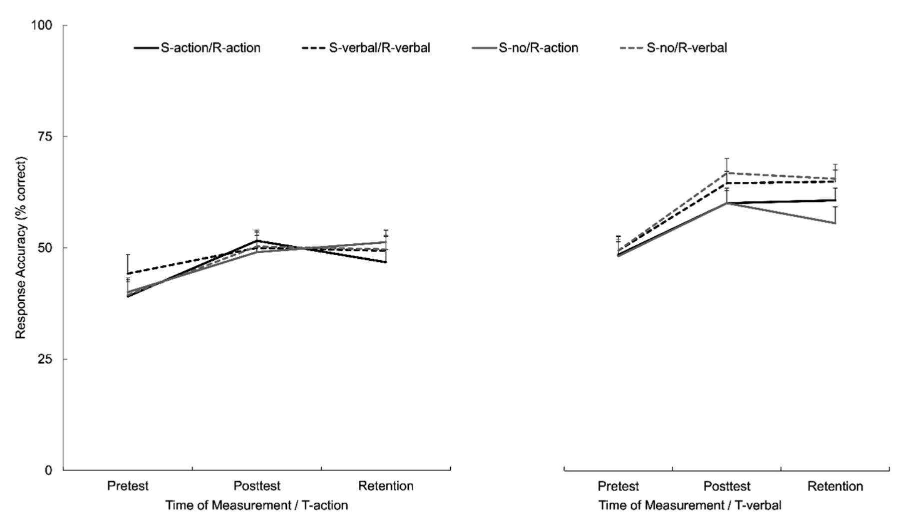

Fig. 5. - Experiment 2: Response accuracy (% correct) as a function of intervention group (S-action/R-action vs. S-verbal/R-verbal vs. S-no/R-action vs. S-no/R-verbal), time of measurement (pretest vs. posttest vs. retention test), and test condition (T-action vs. T-verbal).

The response-initiation analyses (see Table II, Experiment 2) revealed a significant main effect for measurement only, $F(2, 96) = 4.27$, $p < .05$, $\eta_p^2 = .08$. The participants responded earlier in the retention than in the pretest, $t(51) = 2.44$, $p < .05$, $d = 0.36$, without significant differences between pre- and posttest, $t(51) = 1.95$, $p = .06$, $d = 0.27$, as well as between post- and retention test, $t(51) = 0.76$, $p = .45$, $d = 0.09$. The remaining comparisons were insignificant (all *p*s > .66, all $\eta_p^2 < .01$).

Discussion

In Experiment 2, intervention conditions were matched to test conditions by comparing four treatments: (i) with vs. without colour cueing and (ii) with action responses to non-occluded scenes vs. verbal responses to occluded scenes. As experts' gaze paths were presented that had been acquired under the respective test conditions, the crucial comparisons regarded the contrasts between S-action/R-action and S-no/R-action on the one and be-

TABLE II

Relative movement-initiation time (ms) as a function of intervention groups (S-action/R-action vs. S-Verbal/R-Verbal Vs. S-No/R-Action Vs. S-No/R-Verbal) And Time Of Measurement (Pretest Vs. Posttest Vs. Retention Test). The Value Of 0 Ms Denotes The Moment When The Attacker Hits The Ball.

	Group	**Movement-Initiation Time (ms)**
Pretest	S-action/R-action	-281.6 (± 244.2)
	S-verbal/R-verbal	-281.8 (± 260.4)
	S-no/R-action	-247.1 (± 142.1)
	S-no/R-verbal	-206.8 (± 193.4)
Posttest	S-action/R-action	-395.4 (± 128.6)
	S-verbal/R-verbal	-336.8 (± 187.1)
	S-no/R-action	-306.3 (± 69.5)
	S-no/R-verbal	-352.3 (± 265.5)
Retention	S-action/R-action	-367.3 (± 137.2)
	S-verbal/R-verbal	-394.1 (± 169.5)
	S-no/R-action	-324.0 (± 134.9)
	S-no/R-verbal	-349.3 (± 211.7)

tween S-verbal/R-verbal and S-no/R-verbal on the other hand. Advantages of the colour-cue groups could be expected in the respective test condition at least (i.e., T-action and T-verbal), in comparison to Experiment 1, also because the intervention duration was doubled.

The findings show that decision making was generally improved from pre- to post- and retention test regarding response accuracy and response initiation in the T-action test. However, no learning enhancement was found as a function of colour cueing. For gaze behaviour, as in Experiment 1, participants showed a test-specific behaviour with gaze paths closer to the experts' gaze paths if the current test conditions matched the experts' ones. However, gaze behaviour was not differentially affected by the experimental interventions.

Hence, the results further corroborate the findings revealed in Experiment 1. Even if (i) interventions and test conditions are perfectly matched and (ii) the intervention is substantially prolonged, the hypothesis of positive effects of colour cueing on decision-making enhancement cannot be confirmed.

General Discussion

The results of the two reported experiments show that colour-cue gaze-path learning methods do not promote the acquisition of expert-like gaze behaviour or result in improved decision-making (see also Abernethy et al., 2012; Klostermann et al., 2015). Neither adapting training and test conditions to the gaze behaviour to be learned (Experiment 1) nor further improving the intervention-test match by introducing corresponding response conditions or substantially prolonging the intervention phase (Experiment 2) showed an advantage of colour cueing over the simple presentation of untreated training videos. Although colour cueing had no adverting effects, in terms of efficiency, one should refrain from applying gaze-path cueing methods in perceptual training.

An explanation for missing positive effects of gaze-path cueing refers to the speculation that the participants denied to adopt experts' gaze behaviour as they require different information for successful decision making. In this regard, Williams et al. (2009) found that skilled tennis players showed detrimental performance if the virtual opponents' proximal or distal kinematic features were manipulated, whilst less skilled tennis players' performance was impaired in the distal-manipulation condition only. Similarly, Bourne et al. (2013) demonstrated for handball throwing that the neutralisation of proximal information sources selectively affected the anticipation performance of skilled when compared to less-skilled players. As these findings suggest different information extraction strategies as a function of expertise, adopting experts' gaze paths might not result in the desired learning-enhancing effects. Alternatively, missing effects of gaze-path cueing could be explained by the hypothesis suggested by Abernethy et al. (2012) that the coloured patches distract participants' visual attention from the main task of decision-making by drawing their attentional focus away from the locations where the crucial information can be found best. As their empirical study on handball throwing indicated increased searching for other – possibly less functional – kinematic cues in the colour-cue group, the distraction hypothesis of missing positive effects of gaze-path cueing definitely deserves attention in the context at hand. Finally, Cañal-Bruland (2009) suggested that by applying the cueing paradigm (Posner, 1980), lower-order cognitive skills, like signal detection, can be facilitated. However, for higher-order cognitive processes required in complex decision making in sport-specific situations, like selecting and executing respective responses, the cueing paradigm might have no beneficial effect due to low attentional resources. Moreover, from an ecological perspective, the current null findings align with concepts like self-organisation (e.g., Glazier & Robins, 2013) and the constraints-led approach (e.g., Renshaw & Chow, 2019). Instead of explicitly prescribing an ideal gaze pattern, the constraints-led approach suggests enhancing exploration of solutions in the performer-environment system and exploiting self-organisation tendencies of the system. Therefore, by applying visual stimuli to guide gaze behaviour, movement control was prescribed, thus, disabling the self-organisation of the movement system. Because of this dysfunctionality, one should neither expect the system to adopt this behaviour nor anticipate improved decision-making performance.

Instead of taking obtained results for granted, especially in the case of null findings, potential measuring inaccuracies must also be considered. In this vein, our data acquisition could have been contaminated by a random measurement error such that the findings would reflect measurement noise. However, three facts speak against this alternative explanation, namely, that congruent results were found in a previous study (Klostermann et al., 2015), that we applied a highly accurate and reliable gaze-analysis system (cf. Kredel et al., 2015), and that distinct differences in gaze behaviour as a function of test demands (action vs. verbal) were reliably revealed. Hence, little room is left for the interpretation that our data-acquisition and analysis procedures are substantially corrupted by noise.

A further limitation might reside in the selection of our participants. For example, experienced beach-volleyball players might have been better able to adopt the expert gaze pattern due to their previous experience with the task and specific environmental features. This argument is grounded in the general discussion on universal optimal gaze patterns (e.g., Renshaw et al., 2019). It is argued that functional variances inherent in the human movement system are ignored when applying “one-size-fits-all” approaches in motor learning. Consequently, depending on, e.g., movement experience, specific features in training must be addressed to consider, e.g., organismic constraints. Interestingly, however, the most crucial aspect in our gaze training, i.e., the gaze jump to the future position of ball-hand contact, was found in the pretest already. Thus, to improve the anticipation of unknown behaviour, participants already used this specific gaze behaviour. But this coupling was likely developed in different initial situations, and participants could not use this in the current task specificity, i.e., the perception could not be coupled to the appropriate action. However, to disentangle this alternative interpretation, further research is required to, for example, test the effectiveness of the current gaze training in different expertise levels.

Besides the missing confirmation of the expectation of learning-enhancing colour-cueing effects, the last-mentioned finding of distinct gaze differences as a function of test demands deserves a closer look. Whilst the participants (of all groups in both experiments) shifted their focus toward the attacker in the verbal test condition, they preferred a focus closer to the location where they expected the ball to be hit in the action test condition. Notably, such a – in the classification proposed by Vater et al. (2020) – “gaze-anchoring behaviour” was shown not only by the colour-cue groups after intervention (i.e., after learning exactly this gaze behaviour from the repeatedly presented experts’ gaze paths) but also by the control groups and for all groups even in the pretest. Albeit this effect of anchoring gaze closer to the ball if an active response is required instead of a verbal judgement (Dicks et al., 2010), to the best of our knowledge, this is the first study to reveal these task-dependent gaze strategies in novices. Thus, this effect does not – or at least not exclusively – seem to be bound to domain-specific knowledge acquired in the course of practice but should be regarded as a fundamental phenomenon of perception-action coupling (Dicks et al., 2010).

Concluding, based on our results, it is not advisable to apply colour-cueing techniques in sports practice because (i) the potential gain seems to be somewhat limited, especially when considering the extra amount of effort that is required for putting a respective intervention into practice, and (ii) in the worst case, learning is hampered rather than enhanced, may it be due to the acquisition of gaze paths that do not fit the learner’s expertise level (Williams et al., 2009) or due to distracting the learner’s attention from the main task of decision-making (Abernethy et al., 2012). From a more theoretical perspective, the question of whether missing – or even detrimental – effects of gaze-path cueing should be better ascribed to the dysfunctionality of highlighted information or attentional distraction from decision-making would be worthwhile to be pursued in future research, probably best by independently manipulating both factors in empirical studies.

REFERENCES

Abernethy, B., Schorer, J., Jackson, R. C., & Hagemann, N. (2012). Perceptual training methods compared: The relative efficacy of different approaches to enhancing sport-specific anticipation. *Journal of Experimental Psychology: Applied*, *18*, 143-153.

Bourne, M., Bennett, S. J., Hayes, S. J., Smeeton, N. J., & Williams, A. M. (2013). Information underpinning anticipation of goal-directed throwing. *Attention, Perception, & Psychophysics*, *75*, 1559-1569.

Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., & Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. Psychological bulletin, 145(10), 980.

Broadbent, D. P., Causer, J., Williams, A. M., & Ford, P. R. (2015). Perceptual-cognitive skill training and its transfer to expert performance in the field: Future research directions. European journal of sport science, 15(4), 322-331.

Cañal Bruland, R. (2009). Visual cueing in sport specific decision making. *International Journal of Sport and Exercise Psychology*, *7*, 450-464.

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, & Psychophysics*, *72*, 706-720.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*, 523-552.

Glazier, P. S., & Robins, M. T. (2013). Self-organisation and constraints in sports performance. In Routledge handbook of sports performance analysis (pp. 42-51). Routledge.

Hagemann, N., Strauß, B., & Cañal-Bruland, R. (2006). Training perceptual skill by orienting visual attention. *Journal of Sport & Exercise Psychology*, *28*, 143-158.

Klostermann, A., Vater, C., Kredel, R., & Hossner, E. J. (2015). Perceptual training in beach volleyball defence: different effects of gaze-path cueing on gaze and decision-making. *Frontiers in Psychology*, *6:1834*. doi: 10.3389/fpsyg.2015.01834.

Larkin, P., Mesagno, C., Spittle, M., & Berry, J. (2015). An evaluation of video-based training programs for perceptual-cognitive skill development. A systematic review of current sport-based knowledge. *International Journal of Sport Psychology*, *46*, 555-586.

Mann, D. T. Y., Williams, A. M., Ward, P., & Janelle, C. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport & Exercise Psychology*, *29*, 457-478.

Marteniuk, R. G. (1976). Cognitive information processes in motor short-term memory and movement production. In G. E. Stelmach (Ed.), Motor control (pp. 175-186). San Diego, CA: Academic Press.

Prieto, T. E., Myklebust, J. B., Hoffmann, R. G., Lovett, E. G., & Myklebust, B. M. (1996). Measures of postural steadiness: differences between healthy young and elderly adults. *IEEE Transactions on Biomedical Engineering*, *43*, 956-966.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.

Renshaw, I., & Chow, J. Y. (2019). A constraint-led approach to sport and physical education pedagogy. *Physical Education and Sport Pedagogy*, *24*(2), 103-116.

Ryu, D., Kim, S., Abernethy, B., & Mann, D. L. (2013). Guiding attention aids the acquisition of anticipatory skill in novice soccer goalkeepers. *Research Quarterly for Exercise and Sport*, *84*, 252-262.

Savelsbergh, G. J. P., Van Gastel, P., & Van Kampen, P. (2010). Anticipation of a penalty kicking direction can be improved by directing attention through perceptual learning. *International Journal of Sport Psychology*, *41*, 24-41.

Vater, C., Gray, R., & Holcombe, A. O. (2021). A critical systematic review of the Neurotracker perceptual-cognitive training tool. Psychonomic bulletin & review, 28(5), 1458-1483.

Vater, C., Williams, A. M. & Hossner, E.-J. (2020). What do we see out of the corner of our eye? The role of visual pivots and gaze anchors in sport. *International Review of Sport and Exercise Psychology*, *13*(1), 81-103.

Williams, A. M., Ford, P. R., Eccles, D. W., & Ward, P. (2011). Perceptual cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Applied Cognitive Psychology*, *25*, 432-442.

Williams, A. M., Huys, R., Cañal-Bruland, R., & Hagemann, N. (2009). The dynamical information underpinning anticipation skill. *Human Movement Science*, *28*, 362-370.

Manuscript submitted August 2022. Accepted for publication March 2023.