



Prediction of action outcome: Effects of available information about body structure

Johannes Kurz¹ · Fabian Helm² · Nikolaus F. Troje³ · Jörn Munzert¹

Published online: 3 December 2019
© The Psychonomic Society, Inc. 2019

Abstract

Correctly perceiving the movements of opponents is essential in everyday life as well as in many sports. Several studies have shown a better prediction performance for detailed stimuli compared to point-light displays (PLDs). However, it remains unclear whether differences in prediction performance result from explicit information about articulation or from information about body shape. We therefore presented three different types of stimuli (PLDs, stick figures, and skinned avatars) with different amounts of available information of soccer players' run-ups. Stimulus presentation was faded out at ball contact. Participants had to react to the perceived shot direction with a full-body movement. Results showed no differences for time to virtual ball contact between presentation modes. However, prediction performance was significantly better for avatars and stick figures compared to PLDs, but did not differ between avatars and stick figures, suggesting that explicit information about the articulation of the major joints is mainly relevant for better prediction performance, and plays a larger role than detailed information about body shape. We also tracked eye movements and found that gaze behavior for avatars differed from those for PLDs and stick figures, with no significant differences between PLDs and stick figures. This effect was due to more and longer fixations on the head when avatars were presented.

Keywords Prediction · Kinematic information · Soccer penalty · Gaze behavior · Structural body information

Introduction

Perceiving and anticipating the movements of other people is essential to maintain social life and communication with them. Humans use such information, for example, to avoid collisions with other people in crowds or to interpret gestures in communication. These processes also play an essential role in sport tasks in tennis, team handball, or soccer, for example. Sport tasks may serve as central paradigms for anticipation, because athletes often have to react under very limited time constraints. Particularly in interactive sports, athletes have to anticipate the movements of an opponent and the subsequent

trajectory of an object that has been shot by the opponent (e.g., a ball). Early prediction of the outcome of their opponent's movement increases their chance to respond optimally. Athletes gain a competitive advantage if they pick up significant visual information from their opponent's movements that enables them to select an appropriate motor response.

Perception of human motion with an emphasis on kinematic aspects was first studied using the point-light technique. In 1973, Johansson showed that about 10–12 reflecting markers representing the major joints of a human are enough to identify different human actions (e.g., walking, running, or dancing) (Johansson, 1973; see Blake & Shiffrar, 2007, for a review). Point-light displays (PLDs) have been shown to contain sufficient information to identify individuals (Troje, Westhoff, & Lavrov, 2005), to discriminate gender (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Troje, 2002, 2008), to recognize emotions from movement kinematics (Atkinson, Dittrich, Gemmell, & Young, 2004; Dittrich, Troscianko, Lea, & Morgan, 1996), and to identify the actions of a person (Hohmann, Troje, Olmos, & Munzert, 2011; Munzert, Hohmann, & Hossner, 2010).

PLDs, as originally conceived by Johansson, are specifically helpful for the study of research questions related to the effects of

✉ Johannes Kurz
johannes.kurz@sport.uni-giessen.de

¹ NemoLab, Department of Psychology and Sports Sciences, Justus Liebig University, Giessen, Germany

² Department of Psychology and Sports Sciences, Goethe University, Frankfurt/Main, Germany

³ Biomotionlab, Centre for Vision Research and Department of Biology, York University, Toronto, Canada

perceptual organization (structure-from-motion; see Troje, 2013, for details). However, PLDs have also been used to identify the influence of movement kinematics on perception during anticipation in sports (e.g., Abernethy, Gill, Parks, & Packer, 2001; Ward, Williams, & Bennett, 2002). One reason why PLDs were used in these studies was that this technique addressed a technical problem: to isolate kinematic information from other sources of information (e.g., facial expression, hair, or clothes). However, PLDs also provide information about the general proportions of a body's configuration. When research questions are related to the semantic content of the motion (e.g., to predict the shot direction of a tennis serve), the use of stick figures might be a more attractive option. It has been argued that intrinsic movement features can be derived from a single frame in this condition (see Troje, 2013, for details). Stick figures also isolate kinematic information from other sources of information, but they additionally provide explicit information about articulation of the joints without requiring the observer's visual system to infer it from single dots. Thus, stick figures do not confound the task to organize individual markers into an articulated structure (as it is the case for PLDs) with the task to analyze the semantic content of the motion. In addition to PLDs and stick figures, there exist other presentation modes, for instance video clips and computer graphic (CG) animations. These presentation modes depict not only the kinematic information and the explicit articulation of the joints, but also provide further information about a person's body shape. Therefore, the sequence of PLDs, stick figures, and CG animations/video clips reflects an increase of information about body structure that is available to an observer. The question arises whether the availability of more detailed information about body structure also enhances anticipation performance in sports. The existing literature does not provide a clear answer to that question. Several studies did not find significant differences for anticipation performance between PLDs and video clips/CG animations (Fukuhara, Ida, Ogata, Ishii, & Higuchi, 2017; Shim, Carlton, Chow, & Chae, 2005; Shim, Carlton, & Kwon, 2006; Vignais et al., 2009), whereas a number of studies revealed significant differences between these presentation modes (Abernethy et al., 2001; Shim et al., 2005; Ward et al., 2002).

It is not clear why some studies found a difference in performance on video clips and CG animations over PLDs. The fact that some studies see clear differences implies that the power of the experimental design may not have been sufficient in the other studies. It is also not clear where these differences come from. Specifically, it remains unclear whether the observed differences are due to the fact that video clips and CG animations, in contrast to PLDs, provide explicit information about the articulation of the major joints or whether it is because they contain additional, more detailed information about a person's body.

Against this background, the aim of the current study was to investigate whether the previously observed advantage in predicting the action outcome of detailed stimuli over PLDs is

due to the lack of explicit information about the articulation of the major joints in PLDs, or whether it is due to the lack of available information about a person's body shape in PLDs. Up to now, this has not been systematically investigated. We therefore presented three different types of stimuli (PLDs, stick figures, and skinned avatars) with different amounts of available information about body structure of soccer players' run-ups on a large screen and requested participants to predict shot direction (left vs. right) with a full-body movement. Stick figures provide explicit information about the articulation of the body but are still deprived of information about facial identity and detailed body shape.

If information about body shape supports prediction performance, accuracy for avatars should be superior to that of PLDs and stick figures, whereas accuracy for PLDs and stick figures should not differ. If explicit information about articulation of the major joints is the reason for better prediction performance, accuracy for avatars and stick figures should not differ and accuracy for both presentation modes should be superior to PLDs. Of course, both aspects could play a role, which would then place performance in response to stick figures somewhere between the other two presentation modes.

In addition to assessing the accuracy of the observers' responses, we also measured their time to virtual ball contact with the intention of screening our data for possible speed-accuracy trade-offs. We also measured gaze behavior, hoping that potential differences in performance are reflected in differences in the gaze behavior. As it has been shown previously, gaze behavior differs significantly between artificial and more representative experimental conditions (e.g., Dicks, Button, & Davids, 2010; see Kurz & Munzert, 2018, for a review). It also differs depending on the observer's task (Saunders, Williamson, & Troje, 2010). Thus, we expected differences between the different presentation modes for gaze behavior. In more specific terms, we expected significant differences especially between avatars (more representative) and PLDs (more artificial), as shown by Ward et al. (2002). We considered that it is possible that gaze behavior for stick figures could be either more similar to PLDs because in both presentation modes markers represent the major joints and they do not provide information about body shape, or that it could be more similar to avatars because both presentation modes contain explicit information about joint articulation.

Method

Participants

Observers in this study ($N = 13$) were competitive elite (Swann, Moran, & Piggott, 2015) soccer players (age: $M = 21.0$ years, $SD = 2.7$) including goalkeepers, defensive players, midfield players, and forwards. They showed an

average playing experience of 14.8 years ($SD = 4.4$) on a competitive level and reported practicing for an average of 6.2 h per week ($SD = 2.9$). Ten observers were self-declared right-footers; three were left-footers. All had normal or corrected-to-normal vision and were naïve to the aim of the study. Before the experiment started, observers gave written informed consent. The study conformed to the guidelines of the Declaration of Helsinki and was approved by the local Ethics Committee of the Department of Psychology and Sports Sciences at Justus Liebig University Giessen.

Stimulus production

Motion capturing was used to record action performance of ten different right-footed soccer players ($M = 24.1$ years, $SD = 7.4$) who played soccer on a competitive level for an average of 16.6 years ($SD = 9.3$). Motion data were recorded indoors while players shot a standard sized 5 ball (Nike Team Training, SC1911–880) at a target ($\varnothing = 22$ cm) within an indoor goal (3.50×1.75 m). The initial distance between the ball and the goal was 5.25 m. The target was placed in the upper left and in the upper right corner of the goal. For each target, the distance between the center of the target and the goal post and between the center of the target and the crossbar was 40 cm. While performing the task, soccer players were free to choose the number of steps and the angle of the run-up, but were asked to make at least three steps before ball contact.

Kinematic data were recorded by means of an optical motion capture system (Vicon, Oxford, UK) equipped with ten high-speed cameras. The motion-capture system tracked three-dimensional trajectories of retroreflective markers with a spatial accuracy of 1 mm at a sampling rate of 200 Hz. During the task, a set of 41 markers (standard full-body marker set) was attached to the soccer players who wore tight neoprene shirts and shorts.

Most markers were attached directly to the skin or to the tight neoprene. Others, such as those for the head, were attached to an elastic band and the ones on the feet were taped to the soccer players' shoes. Additionally, we attached two markers on opposite sides of the ball to determine its center. Data were preprocessed with Nexus 1.8.5 (Vicon, Oxford, England). The motion capture sequence was clipped such that it started during the third last step at the time when the velocity of the marker placed on the heel of the left foot exceeded a threshold of 10 mm/s. It ended when the velocity of the ball exceeded a threshold of 10 mm/s. The duration of a single stimulus was 1.52 s on average ($SD = 0.26$ s).

The surface marker information from each trial was then used to generate three different presentation modes (Fig. 1): (a) *point-light displays* (PLDs), (b) *stick figures*, and (c) *skinned avatars*. For the PLDs and stick figures, the location of 15 “virtual” markers (Troje, 2002, 2008) positioned at major joints (center of the head, sternum, shoulders, elbows, wrists, center of the pelvis, hips, knees, and ankles) of the body were calculated using Nexus 1.8.5. Several studies (Diaz, Fajen, & Phillips, 2012; Lees & Owens, 2011; Lopes, Jacobs, Travieso, & Araújo, 2014) have shown that the poses of both the kicking and the supporting foot are significant predictors of shot direction. Therefore, we displayed toe markers in PLDs and in stick figures to provide information about the orientation of the feet, which is also present in avatars. Adding the two toe markers resulted in 17 virtual markers applied for PLDs and stick figures. The avatars were reconstructed by means of the MoSh algorithm (Loper, Mahmood, & Black, 2014) to obtain correlated estimates for the performer's individual body shape and body motion for each individual penalty kick. This process has shown a reconstruction error compared to true shape of less than 1 cm (Loper et al., 2014). The PLDs and the stick figures were visualized

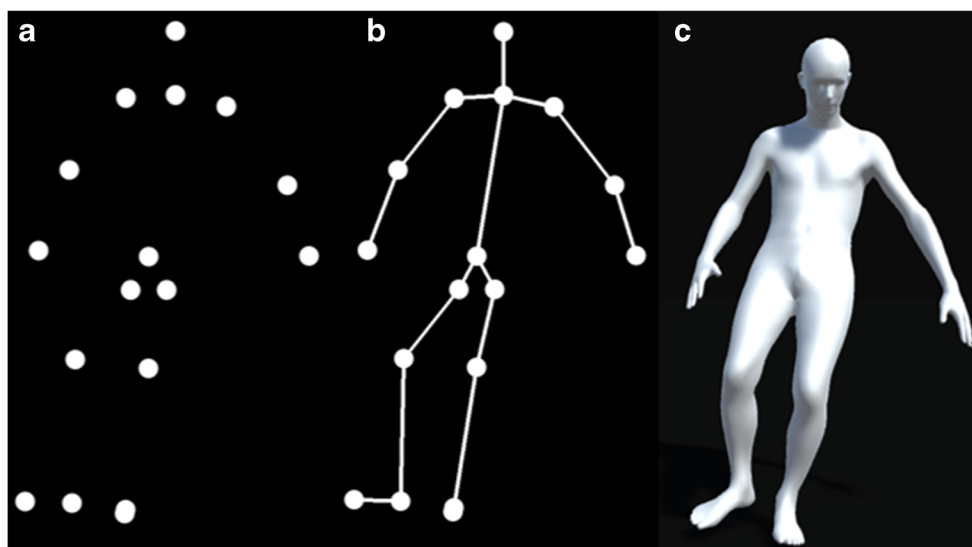


Fig. 1 Amount of available information about body structure of the soccer players in point-light displays (a), stick figures (b), and avatars (c)

using Matlab R2015® (MathWorks, Natick, MA, USA) and the avatars were visualized using Unity3D (Unity Technologies, San Francisco, CA, USA). Data from all presentation modes were down-sampled to a frame rate of 60 Hz (refresh rate of the projector) and rendered into video clips. All stimuli were displayed at veridical speed. Penalty takers were presented from a goalkeeper's perspective.

Apparatus

The stimuli were back-projected (DepthQ HDs3D-1) with a refresh rate of 60 Hz. The distance between the screen and the observers was 3.0 m. Players were depicted with a visual angle between 23° and 25° at ball contact. Stimuli were presented with Matlab R2015® using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Observers stood on a force plate integrated into the floor in front of a large screen (3.2 × 2.1 m) on which the penalty takers were displayed. Observers were required to respond to perceived shot direction (left vs. right) with a full-body movement (with not only a step but also an arm movement in the respective direction). Time to virtual ball contact of the observers was recorded using a force plate (1,000 Hz; Kistler 9281EA, Sindelfingen, Germany) with Nexus 1.8.5 and remote-controlled by Matlab R2015®. Observers' responses were also recorded using the force plate.

Gaze behavior was recorded with a binocular mobile head-mounted eye tracker (SMI, Teltow, Germany) using iViewETG (Version 2.1) recording software. The eye tracker was connected to a mobile recording unit (Samsung Galaxy S4 GT-I9506, Yateley, UK) via a USB that was placed in a belt bag while performing the task. Gaze data were recorded with a frame rate of 60 Hz. After recording, gaze data were exported to a laptop and analyzed frame by frame with BeGaze software (Version 3.5.101). Before starting the experiment, we conducted a three-point calibration that was repeated if necessary. The accuracy of the gaze position (average angle between the actual gaze position and the one measured by the eye tracker) was about 0.5° and the spatial precision (dispersion of recorded gaze points during a fixation) was about 0.1°.

Before starting each trial, an external trigger signal was recorded by the eye tracker and by Matlab R2015® (responsible for stimulus playback and recording ground reaction forces) to synchronize stimuli, gaze data, and ground reaction forces.

Design and procedure

Observers had to predict the shot direction (left vs. right) from presentations of run-ups that were faded out at ball contact. Each participant's reaction was recorded when reacting to three different display mode conditions: (a) PLDs, (b) stick figures, and (c) avatars. Conditions were presented in blocks and they were counterbalanced across observers. The order of the stimuli

was randomized within conditions, but was held constant for each participant. No feedback was provided. Observers were instructed to perform a full-body movement with not only a step but also an arm movement in the respective direction to try to save the ball. Previous studies have shown that more natural response behavior produces more valid information compared to artificial experimental conditions (Dicks, Button, & Davids, 2010; Mann, Abernethy, & Farrow, 2010; see Kurz & Munzert, 2018, for a review). With our instruction, we aimed to create representative experimental conditions with a high degree of perception–action coupling.

Before starting the experiment, observers performed six practice trials – two in each condition. Stimuli presented in practice trials were not used in the experiment. In the main experiment 120 stimuli were presented, 40 stimuli per condition. Half of the 40 stimuli in each condition included shots to the left and half to the right. Each stimulus was presented only once per condition. At the beginning of each trial, observers were instructed to stand still with the right foot placed on a force plate and to direct their gaze toward a fixation cross presented on the screen. After the fixation cross disappeared, the first frame of the stimulus was presented for 1 s. Then the presentation of the run-up started automatically. At ball contact of the presented stimulus, the stimulus disappeared and was replaced by a black screen. Then, 400 ms after the stimulus had disappeared, a sound signal indicated that the virtual ball had hit the goal line. The sound signal was presented to encourage observers to respond as quickly as possible. Inter-trial intervals were set at 2 s.

Data analysis

Accuracy and time to virtual ball contact

Accuracy was defined in terms of the correct response of the observers relative to the actual shot direction. We analyzed the data using signal detection theory and calculated sensitivity (d') and response bias (c). Negative values for response bias were defined as a preference for left-side reactions and positive values indicated right-side preferences. The direction to which observers responded was determined from the force plate data. The following data processing was applied on a trial-by-trial basis. First, ground reaction forces were filtered using a fourth-order low-pass Butterworth filter with a cut-off frequency of 20 Hz. Second, the level of each trial was zeroed by subtracting the mean of the first 200 frames from each frame. Third, the direction of the horizontal ground reaction force was analyzed at the time when it first exceeded 12.5 N.

Time to virtual ball contact was also determined using the force plate. Time to virtual ball contact was defined as the time interval between ball contact and the time at which the ground reaction force first reached 12.5 N. Trials where this happened outside the interval $-400 \text{ ms} < \text{time to virtual ball contact} <$

400 ms were discarded. We chose this lower boundary because a response 400 ms before ball contact gives a soccer player in a real-game situation the opportunity to respond to the goalkeeper's premature reaction and kick the ball to the opposite side (Morya, Ranvaud, & Pinheiro, 2003). We chose the upper boundary because a response 400 ms after ball contact would be too late to prevent the scoring of a goal. In total, 16.5% of the trials were discarded because the time to virtual ball contact was outside the predefined interval.

Eye-tracking data

Percentage viewing time was defined as the percentage of time gaze was directed toward different locations on the screen. We defined five areas of interest to analyze gaze behavior: (1) *head*; (2) *upper body* (which also included the arms); (3) *hip*; (4) *legs* (we did not differentiate between gaze directed to the left and the right leg or foot because it is almost impossible to distinguish clearly between left and right in PLDs); (5) *other* (when the gaze was directed toward any other location). A total of 0.4% of gaze data showed missing data. Percentage viewing times of each area of interest and the missing data add up to 100%. Recording started when the gaze of the participant was first directed away from the fixation cross and toward one of the four areas of interest for the first time: head, upper body, hip, or legs. It ended at the point of ball contact when the presentation was occluded. For specific analyses of percentage viewing times, *other* locations and missing data were excluded. In general, this involved about 1.6% of the total viewing time. Percentage viewing time was separated into three parts: (a) *static*: this part included the perception of the static frame before the motion stimulus started. The depicted movement sequence was divided into two equal parts: (b) *First half* and (c) *Second half*. The three parts of the stimulus were designed to analyze gaze behavior separately when perceiving only structural information (static), dynamic information that does not provide significant information of shot direction (first half), and dynamic information that provides significant information of shot direction (second half).

For sensitivity, we calculated separate one-sample *t*-tests for each condition against the value 0 (chance level). Additionally, sensitivity was analyzed using univariate ANOVAs with repeated measures for the factor display mode (PLDs vs. stick figures vs. avatars). Helmert contrasts were used to test for differences between these levels in order to avoid Bonferroni adjustments (Perneger, 1998). Response bias was also analyzed using separate one-sample *t*-tests for each condition against the value 0. Time to virtual ball contact was analyzed using a 2 (response: correct vs. incorrect) \times 3 (mode: PLDs vs. stick figures vs. avatars) ANOVA with repeated measures on both factors. Gaze behavior in terms of the percentage viewing time for all three time periods (static / first half / second half) was analyzed separately using 3 (mode:

PLDs vs. stick figures vs. avatars) \times 4 (area: head vs. upper body vs. hip vs. legs) ANOVAs with repeated measures on both factors. Post hoc comparisons for the ANOVA analyzing gaze behavior were calculated using *t*-tests with Bonferroni corrections; effect sizes were calculated for the *t*-tests as Cohen's *d* and for the ANOVAs as partial eta squared; and the significance level was set at .05.

Results

Accuracy

On average, observers responded correctly in 68.2% of the cases. Average sensitivity measured as *d'* was .98. Broken down into display conditions (Fig. 2, Table 1) revealed significant sensitivity for PLDs, $t(12) = 3.34, p < .01, d = 0.93$, stick figures, $t(12) = 5.31, p < .001, d = 1.47$, and avatars, $t(12) = 5.54, p < .001, d = 1.54$. These findings indicate that observers are capable of recognizing shooting direction irrespective of condition.

Results on sensitivity (Table 1) between presentation modes revealed a significant main effect of mode, $F(2, 24) = 7.72, p < .01, \eta_p^2 = .39$. Planned Helmert contrasts revealed that sensitivity for PLDs was significantly worse compared to the other two conditions (stick figures and avatars), $t(12) = 3.68, p < .01$, but that stick figures and avatars did not differ significantly, $t(12) = 1.39, p = .18$. Response bias (Table 1) was significantly different from zero for PLDs, $t(12) = 2.53, p < .05$, and for stick figures, $t(12) = 2.93, p < .05$. Observers preferred to respond to the left side in these conditions.

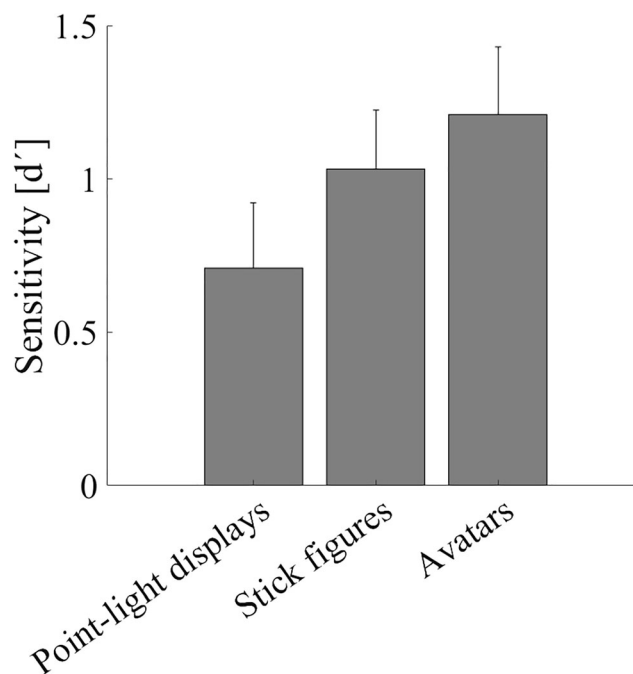


Fig. 2 Sensitivity scores (and standard errors) in terms of *d* prime (*d'*) for point-light displays, stick figures, and avatars

Table 1 Mean scores (and SEMs) for accuracy, sensitivity (d'), response bias (c), and time to virtual ball contact for correct and incorrect trials. Note that negative biases indicate preference for the left side

Representation	Sensitivity (d')	Response bias (c)	RT correct (ms)	RT incorrect (ms)
PLDs	0.71 (0.21)	-0.43 (0.18)	-166.9 (42.9)	- 220.0 (50.0)
Stick figures	1.03 (0.24)	-0.21 (0.08)	-180.4 (36.8)	-204.2 (47.9)
Avatars	1.21 (0.24)	-0.07 (0.11)	-200.5 (34.8)	- 259.9 (49.2)

PLDs point-light displays, *TTB* time to virtual ball contact

Response biases for avatars, $t(12) = .67$, *ns*, were not different from zero.

Time to virtual ball contact

Results on time to virtual ball contact (Table 1) revealed a significant main effect of response (correct vs. incorrect), $F(1,12) = 5.62$, $p < .05$, $\eta_p^2 = .34$. Times to ball contact were on average 45 ms later for correct responses than for incorrect responses. The main effect of mode, $F(2,24) = 1.88$, $p = .18$, and the Response \times Mode interaction, $F(3,42) = .70$, *ns*, did not attain significance.

Percentage viewing time

Percentage viewing time for different areas of interest and for different presentation modes is shown in Fig. 3. The results of respective ANOVAs are presented in Table 2. The interesting results that can be obtained from these are the following: (1) In the static part, results show, irrespective of conditions, that gaze was directed for a similar duration of time toward the upper body, the hip, and the legs. When avatars were presented, gaze was also directed toward the head during the static

part. (2) In contrast, in the first and the second half of the run up, gaze was directed mainly toward the legs irrespective of condition. (3) Results also revealed that in the static part and the first half of the run up observers directed their gaze for significantly longer toward the head and for significantly shorter toward the legs for avatars compared to PLDs and stick figures (both $p < .05$). (4) However, in the second half of the run-up observers showed similar gaze behavior across presentation modes.

Discussion

In the present study, we aimed to investigate whether the advantage of detailed stimuli over PLDs in prediction performance is due to explicit information about the articulation of the major joints in PLDs, whether it is due to the additional amount of available information about a person’s body shape, or whether it is due to the combination of the respective features. To get a better understanding of why the presentation modes differ concerning prediction performance we also investigated time to virtual ball contact as well as gaze behavior in terms of percentage viewing time.

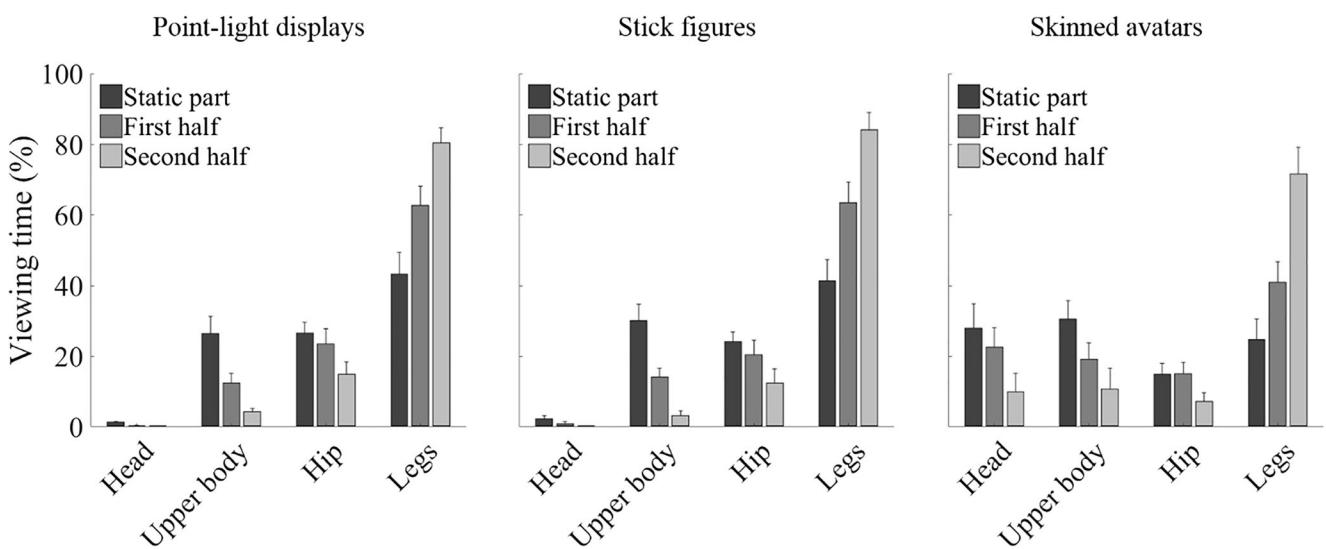


Fig. 3 Percentage viewing time (means and standard error) on the four areas of interest: head, upper body, hip, and legs, for the three presentation modes (point-light displays, stick figures, and avatars) for each part of the stimulus presentation (static part, first half, and second half)

Table 2 Summary of ANOVA results on percentage viewing times at different areas of interest

Part	Source	df	F	p	η_p^2
Static	Rep	2, 24	.33	.720	.03
	Area	3, 36	6.51	< .050	.35
	Rep \times Area	6, 72	8.14	< .001	.40
First half	Rep	2, 24	.75	.490	.06
	Area	3, 36	33.54	< .001	.74
	Rep \times Area	6, 72	8.04	< .001	.40
Second half	Rep	2, 24	.47	.630	.04
	Area	3, 36	102.49	< .001	.90
	Rep \times Area	6, 72	2.46	.100	.17

Rep representation, Area area of interest

Accuracy

Our results on predicting shot direction revealed significant differences between avatars and PLDs and between stick figures and PLDs, with higher sensitivity for avatars and stick figures compared to PLDs resulting in higher sensitivity. Avatars and stick figures did not differ significantly. Thus, stick figures seem to provide similar information to avatars, but not to PLDs. Our findings therefore suggest that the differences in prediction performance between PLDs and video clips that have been observed previously (Abernethy et al., 2001; Shim et al., 2005; Ward et al., 2002) result mainly from the lack of explicit articulation of the major joints in PLDs. The only difference between PLDs and stick figures is that the joints are connected by lines, which helps to identify the specific body configuration for biological motion. These features have to be reconstructed for PLDs. On the whole, this functions well for PLDs, but the reconstruction process requires additional processing capacities and is potentially error prone.

One major result of the present study is that we found significant differences between PLDs and the other two modes, but no relevant differences between stick figures and avatars. We have to carefully interpret the latter result, because results for avatars exceeded those for stick figures on a descriptive level with a level of significance of $p = .18$ (see also Fig. 2). However, we would argue that the main differences between the three presentation modes are due to reconstruction processes as outlined in the previous paragraph. The descriptive advantage of avatars over stick figures, which can be traced back to a better depiction of the bodily contours, seems to be too small to produce significant effects. Therefore, information about body shape seems to be negligible.

Our results replicate and extend the work of Abernethy et al. (2001), Ward et al. (2002), and Shim et al. (2005), who found differences between video clips and PLDs. However, our results differ from those of Fukuhara et al. (2017) and Vignais et al. (2009), who did not find significant differences

between CG animations and PLDs. Our results do not provide insight into why Fukuhara (2017) and Vignais et al. (2009) were not able to detect these differences. However, the technique of creating avatars in the present study (Loper et al., 2014) included a more sophisticated level of stimulus construction compared to previous studies.

Time to virtual ball contact

For time to virtual ball contact, we found the expected difference between correct and incorrect responses that is typical for psychophysical experiments. We did not observe differences in times to ball contact as a function of presentation mode, though. That result gives us more confidence into the measured effects of display mode on sensitivity as we can assume that they were not confounded by speed-accuracy trade-off effects. The finding also supports the results reported by Vignais et al. (2009) that observers' temporal responses are similar when stimuli that contain different amounts of available information about body structure were presented. Several studies (Diaz et al., 2012; Lees & Owens, 2011; Lopes et al., 2014) have shown that significant information about the shot direction first becomes available at about 200–250 ms before ball contact. Early responses are therefore more prone to guessing. Additionally, our results revealed that when observers responded correctly, they responded significantly later compared to when they respond incorrectly, irrespective of the presentation mode.

Gaze behavior

Gaze behavior in terms of percentage viewing time was split into three parts (static, first half, and second half of the run-up). In the first static part, results show, irrespective of conditions, that gaze was directed for similar durations toward the upper body, the hip, and the legs. This result is somewhat surprising, because stick figures and avatars provide in the static condition more information about the body configuration compared to PLDs that should have evoked a different gaze pattern. However, results revealed a similar gaze pattern across presentation modes. This result may be due to the presentation type of the stimuli. Presentations were very similar across trials and conditions, always starting with an opponent depicted in a frontal view, who always started his movement in the same direction. One obvious possibility is that observers could generalize knowledge learned during the dynamic part of the stimulus presentation and transfer this knowledge to the static part of the stimulus presentation. When avatars were presented, gaze was also directed toward the head during the static part. In contrast, in the second and the third parts (which represent the run-up of the penalty taker), gaze was directed mainly toward the legs

irrespective of condition. These results reveal that gaze behavior is directed primarily toward task-relevant locations (Hayhoe & Ballard, 2005). In the second part and in particular in the third part of stimulus presentation, observers focused gaze on information-rich areas that have been identified as such in kinematic analyses (Diaz et al., 2012; Lees & Owens, 2011; Saunders, Williamson, & Troje, 2010). This gaze behavior is described as “pro-active” (Hayhoe & Ballard, 2005; Kurz, Hegele, & Munzert, 2018) because gaze is directed toward a location at which an event is expected – in our case, ball contact.

As expected, our results showed that gaze behavior in terms of the percentage viewing time differed between presentation modes. These differences were found in the first and the second part, but not in the third part. However, percentage viewing time did not differ between PLDs and stick figures in all three parts. When avatars were presented, observers directed their gaze for a small but significant time span toward the head, in the first and the second parts. This is in accordance with previous studies (e.g., Dicks et al., 2010; Savelsbergh, van der Kamp, Williams, & Ward, 2002; Savelsbergh, Van der Kamp, Williams, & Ward, 2005) that used either video clips or live, real-world soccer players for stimulus presentation. In contrast, when PLDs and stick figures are presented, observers hardly ever directed their gaze toward the head irrespective of the time period. A feasible explanation for this gaze pattern is that the head in PLDs and stick figures is presented by only one point representing its center. Observers probably seek information about head orientation, eye gaze, and facial expression, which are all absent in both PLDs and stick figures.

Conclusion

Up to now, it has remained unclear whether the advantage of detailed stimuli over PLDs is due to the lack of explicit information about articulation of the major joints or the lack of available information about a person’s body shape in PLDs. Our results suggest that the explicit information about articulation of the major joints (which is present in stick figures and avatars) is mainly relevant for improvement of the prediction performance in penalty kicks. Information about body shape (which is only present in avatars) seems to improve prediction performance only slightly.

Gaze behavior is affected by information about body shape (avatars). These differences between avatars and the presentation modes of PLDs and stick figures were only present at the static part of stimulus presentation and not during the movement sequence. Therefore, we suggest that gaze behavior did not affect prediction performance.

Data availability None of the data or materials for the experiments reported here are available, and none of the experiments were preregistered.

References

- Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception*, 30(2), 233–252. <https://doi.org/10.1068/p2872>
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717–746. <https://doi.org/10.1068/p5096>
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, 58, 47–73. <https://doi.org/10.1146/annurev.psych.57.102904.190152>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Diaz, G. J., Fajen, B. R., & Phillips, F. (2012). Anticipation from biological motion: The goalkeeper problem. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 848–864. <https://doi.org/10.1037/a0026962>
- Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behavior under un situ and video simulation task constraints reveals differences in information pickup for perception. *Attention, Perception & Psychophysics*, 72(3), 706–720. <https://doi.org/10.3758/APP>
- Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727–738. <https://doi.org/10.1068/p250727>
- Fukuhara, K., Ida, H., Ogata, T., Ishii, M., & Higuchi, T. (2017). The role of proximal body information on anticipatory judgment in tennis using graphical information richness. *PLOS ONE*, 12(7), 1–11. <https://doi.org/10.1371/journal.pone.0180985>
- Hayhoe, M. M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4), 188–194. <https://doi.org/10.1016/j.tics.2005.02.009>
- Hohmann, T., Troje, N. F., Olmos, A., & Munzert, J. (2011). The influence of motor expertise and motor experience on action and actor recognition. *Journal of Cognitive Psychology*, 23(4), 403–415. <https://doi.org/10.1080/20445911.2011.525504>
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211. <https://doi.org/10.3758/BF03212378>
- Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What’s new in Psychtoolbox-3? *Perception 36 ECV Abstract Supplement*. <https://doi.org/10.1068/v070821>
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580.
- Kurz, J., Hegele, M., & Munzert, J. (2018). Gaze behavior in a natural environment with a task-relevant distractor: How the presence of a goalkeeper distracts the penalty taker. *Frontiers in Psychology: Cognitive Science*, 9:19, 1–14. <https://doi.org/10.3389/fpsyg.2018.00019>
- Kurz, J., & Munzert, J. (2018). How the experimental setting influences representativeness: A review of gaze behavior in football penalty takers. *Frontiers in Psychology: Movement Science and Sport Psychology*, 9:682. <https://doi.org/10.3389/fpsyg.2018.00682>
- Lees, A., & Owens, L. (2011). Early visual cues associated with a directional place kick in soccer. *Sports Biomechanics*, 10(2), 125–134. <https://doi.org/10.1080/14763141.2011.569565>
- Loper, M., Mahmood, N., & Black, M. J. (2014). MoSh: Motion and shape capture from sparse markers. *ACM Transactions on Graphics*, 33(6), 1–13. <https://doi.org/10.1145/2661229.2661273>

- Lopes, J. E., Jacobs, D. M., Travieso, D., & Araújo, D. (2014). Predicting the lateral direction of deceptive and non-deceptive penalty kicks in football from the kinematics of the kicker. *Human Movement Science*, 36, 199–216. <https://doi.org/10.1016/j.humov.2014.04.004>
- Mann, D. L., Abernethy, B., & Farrow, D. (2010). Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks. *Acta Psychologica*, 135(1), 17–23. <https://doi.org/10.1016/j.actpsy.2010.04.006>
- Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society B: Biological Sciences*, 258, 273–279. <https://doi.org/10.1098/rspb.1994.0173>
- Morya, E., Ranvaud, R., & Pinheiro, W. M. (2003). Dynamics of visual feedback in a laboratory simulation of a penalty kick. *Journal of Sports Sciences*, 21(2), 87–95. <https://doi.org/10.1080/0264041031000070840>
- Munzert, J., Hohmann, T., & Hossner, E. (2010). Discriminating throwing distances from point-light displays with masked ball flight. *European Journal of Cognitive Psychology*, 22, 247–264. <https://doi.org/10.1080/09541440902757975>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ (Clinical Research Ed.)*, 316(7139), 1236–1238. <https://doi.org/10.1136/bmj.316.7139.1236>
- Saunders, D. R., Williamson, D. K., & Troje, N. F. (2010). Gaze patterns during perception of direction and gender from biological motion. *Journal of Vision*, 10(11), 9–9. <https://doi.org/10.1167/10.11.9>
- Savelsbergh, G. J. P., van der Kamp, J., Williams, A. M., & Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of Sports Sciences*, 20, 279–287. <https://doi.org/10.1080/00140130500101346>
- Savelsbergh, G. J. P., Van der Kamp, J., Williams, A. M., & Ward, P. (2005). Anticipation and visual search behaviour in expert soccer goalkeepers. *Ergonomics*, 48(11–14), 1686–1697. <https://doi.org/10.1080/00140130500101346>
- Shim, J., Carlton, L. G., Chow, J. W., & Chae, W.-S. (2005). The use of anticipatory visual cues by highly skilled tennis players. *Journal of Motor Behavior*, 37(2), 164–175. <https://doi.org/10.3200/JMBR.37.2.164-175>
- Shim, J., Carlton, L. G., & Kwon, Y. H. (2006). Perception of kinematic characteristics of tennis strokes for anticipating stroke type and direction. *Research Quarterly for Exercise and Sport*, 77(3), 326–339. <https://doi.org/10.1080/02701367.2006.10599367>
- Swann, C., Moran, A., & Piggott, D. (2015). Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychology of Sport and Exercise*, 16(P1), 3–14. <https://doi.org/10.1016/j.psychsport.2014.07.004>
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2(5), 371–387. <https://doi.org/10.1167/2.5.2>
- Troje, N. F. (2008). Retrieving information from human movement patterns. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding Events: How Humans See, Represent, and Act on Events* (pp. 308–334). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195188370.003.0014>
- Troje, N. F. (2013). What is biological motion? Definition, stimuli, and paradigms. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and Interpretation of Animacy, Agency and Intention* (pp. 13–36). MIT Press. <https://doi.org/10.7551/mitpress/9780262019279.003.0002>
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, 67(4), 667–675. <https://doi.org/10.3758/BF03193523>
- Vignais, N., Bideau, B., Craig, C., Brault, S., Multon, F., Delamarche, P., & Kulpa, R. (2009). Does the level of graphical detail of a virtual handball thrower influence a goalkeeper's motor response? *Journal of Sports Science and Medicine*, 8(4), 501–508.
- Ward, P., Williams, A. M., & Bennett, S. J. (2002). Visual search and biological motion perception in tennis. *Research Quarterly for Exercise and Sport*, 73(1), 107–112. <https://doi.org/10.1080/02701367.2002.10608997>

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