Vision Research 112 (2015) 68–82

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Assessing the precision of gaze following using a stereoscopic 3D virtual reality setting

Artin Atabaki, Karolina Marciniak, Peter W. Dicke, Peter Thier *

Hertie-Institute for Clinical Brain Research, Department of Cognitive Neurology, University of Tübingen, Germany

ARTICLE INFO

Article history:
Received 4 September 2014
Received in revised form 9 April 2015
Available online 14 May 2015

Keywords:
Gaze following
Human avatar
Stereoscopic 3D environment
Social attention

ABSTRACT

Despite the ecological importance of gaze following, little is known about the underlying neuronal processes, which allow us to extract gaze direction from the geometric features of the eye and head of a conspecific. In order to understand the neuronal mechanisms underlying this ability, a careful description of the capacity and the limitations of gaze following at the behavioral level is needed. Previous studies of gaze following, which relied on naturalistic settings have the disadvantage of allowing only very limited control of potentially relevant visual features guiding gaze following, such as the contrast of iris and sclera, the shape of the eyelids and – in the case of photographs – they lack depth. Hence, in order to get full control of potentially relevant features we decided to study gaze following of human observers guided by the gaze of a human avatar seen stereoscopically. To this end we established a stereoscopic 3D virtual reality setup, in which we tested human subjects' abilities to detect at which target a human avatar was looking at. Following the gaze of the avatar showed all the features of the gaze following of a natural person, namely a substantial degree of precision associated with a consistent pattern of systematic deviations from the target. Poor stereo vision affected performance surprisingly little (only in certain experimental conditions). Only gaze following guided by targets at larger downward eccentricities exhibited a differential effect of the presence or absence of accompanying movements of the avatar's eyelids and eyebrows.

1. Introduction

The identification of another person's gaze direction is crucial for social interactions as it allows us to shift our focus of attention to the object of interest of our peer, i.e. to establish joint attention (Scaife & Bruner, 1975; Emery, 2000). By associating the object of joint attention with contextual information we may develop insight into the other one's mental state, in other words to develop a theory of mind (Baron-Cohen, 1995; Dalton et al., 2005). Consider for instance a person, staring at a spider. As the face expresses terror, you may easily conclude that this person is deeply afraid of spiders. Human gaze following is largely determined by eye gaze, most probably a direct consequence of the rich landmarks offered by the human eye. Unlike the eyes of other primate species, the human eye offers a dark center (the pupil and the iris), surrounded by a white area, the sclera, offering a rich contrast easily detectable from quite a distance (Kobayashi & Kohshima, 1997) and therefore suitable to support social interactions also in more extended groups. Despite the ecological importance of gaze following, little is known about the underlying neuronal processes, which allow us to extract gaze direction from the geometric features of the eyes and head of a conspecific. In order to address the neuronal mechanisms underlying gaze following, a careful description of the capacity and the limitations at the behavioral level of gaze following is needed.

Previous studies examining the gaze following ability in humans used naturalistic settings, in which the eye gaze direction of a fellow human being ("sender") had to be evaluated (Anstis, Mayhew, & Morley, 1969; Cline, 1967; Gibson & Pick, 1963; Symons et al., 2004). These experiments revealed that the precision of gaze following is highly accurate. Furthermore, they showed a constant overestimation of gaze following along the horizontal axis. However, these studies used only a limited set of stimuli. A more recent study (Bock, Dicke, & Thier, 2008) characterized in more detail the spatial resolution of observers' ability to detect targets on a circle singled out by a human sender. The participants in this study were able to detect a target on a ring (eccentricity 15°) at which the sender was gazing at with a mean tangential error of 2.81° visual angle, and they exhibited a systematic response bias

* Corresponding author at: Hertie-Institute for Clinical Brain Research, Department for Cognitive Neurology, Otfried-Müller-Strasse 27, 72070 Tübingen, Germany.
E-mail address: thier@uni-tuebingen.de (P. Thier).

http://dx.doi.org/10.1016/j.visres.2015.04.015
0042-6989/

for the cardinal-axes (i.e. the vertical and the horizontal axes). As the spatial arrangement of the stimuli, the sender was gazing at was confined to a circle of fixed radius the authors could not address the generality of their observations, i.e. if the comparably high degree of precision and a bias towards the cardinal axes might also hold for target positions outside the circle tested.

Defining the orientation of our peer's eyes relative to objects in the world requires a consideration of the eyes as well as the face (Langton, Watt, & Bruce, 2000). That indeed both matter is convincingly demonstrated by the Wollaston effect (Wollaston, 1824) which is characterized by a shift in the perceived orientation of a sender's eyes, induced by reorienting facial contours. Actually, work by Todorovic (2006) demonstrated that the perceived gaze direction is not the simple sum of gaze and head/facial orientation relative to an external frame, but the result of a calculation, in which the iris eccentricity is given a stronger weight than the head orientation. However, a rigorous assessment of the influence of different facial components and the head orientation has not been carried out yet, a major reason being that our understanding of the features used in order to retrieve eye gaze is still limited. This is to a large extent a consequence of the fact that previous work has largely relied on naturalistic settings or photographs which have the disadvantage of allowing only very limited control of potentially relevant visual features guiding gaze following. The comparative studies by Kobayashi and Kohshima (1997) clearly suggests an important role of the border between the white sclera and the dark center of the human eye. An important role of this contrast has also been suggested by psychophysical experiments by Ando (2002, 2004) who showed that besides the geometric features also the contrast between the iris and the sclera and neighboring facial structures underlie directional judgments. The important role of the position of the dark center is also evinced by experiments, in which subjects were exposed to eyes with inverted gray values (dark sclera and white iris/pupil). In this case, gaze perception was massively impaired (Ricciardelli, Baylis, & Driver, 2000; Sinha, 2000) as observers typically shifted their gaze in the direction of the dark sclera, rather than in the direction of the (bright) pupil. These observations clearly indicate that not only the overall shape of the transition between iris and sclera, i.e. the geometry of the configuration matters, but also luminance values, the breadth of contrast gradients and, last but not least, the reference provided by neighboring facial features (Frischen, Bayliss, & Tipper, 2007).

In order to achieve better control of potentially relevant features we decided to study gaze following of human observers guided by the gaze of a human avatar seen stereoscopically in a 3D virtual reality setup. This setup allowed us to test human subjects' abilities to detect at which target object, presented midway between the observer and the avatar, the avatar was looking at. By comparing the precision of following the avatar's gaze in a number of specific conditions with gaze following in a natural setting (this study, Anstis, Mayhew, & Morley, 1969; Bock, Dicke, & Thier, 2008; Cline, 1967; Gibson & Pick, 1963; Symons et al., 2004), we show that the study of gaze following in a virtual setting reproduces the key features of gaze following in naturalistic settings and, moreover, it indeed provides additional insights into relevant visual features.

2. General methods section

2.1. Setup generating 3D eye gaze avatars

Stereoscopic visual stimulation was realized with 2 projectors (Epson®, 1735 series) each equipped with spectral comb filters (Infitec®), whose spectral pass bands did not overlap. The projectors operated at a frame rate of 60 Hz each and projected onto the same area of a back-projection translucent screen placed at a distance of 100 cm (experiments 1 and 2) and 80 cm (experiments 3 and 5) in front of the subjects tested. Subjects were seated fronto-parallel to the screen with their nasion aligned with the centre of the screen. A chin-rest assured the same height for the avatar's eyes and the observer's eyes. Subjects wore goggles, fitted with comb filter glasses (Infitec® glasses), each tuned to the omission spectrum of one of the two projectors. As a consequence, each projector's image was seen by one eye only. A Linux computer running the universal open source measurement system nrec (see http://nrec.neurologie.uni-tuebingen.de) was used to generate the visual stimuli, to control the course of the experiment and to collect data.

The human avatar was taken from POSER ver. 7. The contrast difference between the iris-sclera of the avatar face was very high, around 70 times, the measurement taken at the gray level difference across the sclera-iris and boundaries. The size of the avatar face in every experiment was adjusted to be comparable to a real face in a comparable natural setting. The iris shifted from the straight ahead position to the final gaze position; hence the gaze shift as mentioned in the discussion was induced by implied motion. The face with its characteristic was the default POSER 7 face. In order to generate a 3D view of the human avatar with his eyes in different orientations relative to the head and target items midway between the avatar and the observer, we used graphics software, based on the OpenGL library, developed in our laboratory (Benz & Hübner, 2008). This program allowed us to render the required views of the avatar, considering the geometry of the triadic constellation and furthermore assuming that the eyes of the observers would have an interocular distance of 60 mm on the side of the observer. Actually, the mean interocular distance in the group of subjects tested was 64.12 mm (SD = Standard deviation of 2.4 mm), a deviation from our assumption, whose relevance will be addressed later in the discussion.

2.2. Participants

The numbers regarding the different participants in the 5 experiments will be provided in the methods section related to the respective experiments. This study fully conforms with the code of ethics of the World Medical Association (Declaration of Helsinki) (see World Medical Association, 1964) and was approved by the Ethics Review Board of the Medical Faculty of Tübingen University.

2.3. Experiments

In experiments 1, 2, 3 and 5 a trial started with the presentation of a 3D image of the human avatar (for example see Figs. 2A, 3A, 4A, 6A) looking straight ahead at the subjects. The avatar was a caucasian mid-thirty male person offered by POSER ver. 7. The human observers perceived spherical target objects (blue circles, diameter of 0.6° visual angle) stereoscopically midway between themselves and the human avatar. The targets' spatial arrangement varied between the experiments and is described below. Superimposed on the avatar's nasion a fixation dot (red circle, diameter of 0.6°, in the depth-plane of the avatar) prompted the human observers to acquire fixation. Furthermore, in all experiments the target objects were placed in a plane that was fronto-parallel relative to both the avatar and themselves. After 1 s the human avatar shifted eye gaze to one of the targets, while keeping head orientation straight ahead during the entire trial. The avatar shifted gaze from one position to the other from one frame to the next, whereas a human sender generates a saccade, which lasts several 10 ms, in other words, having a duration of several frames. We will discuss the possible implications of this

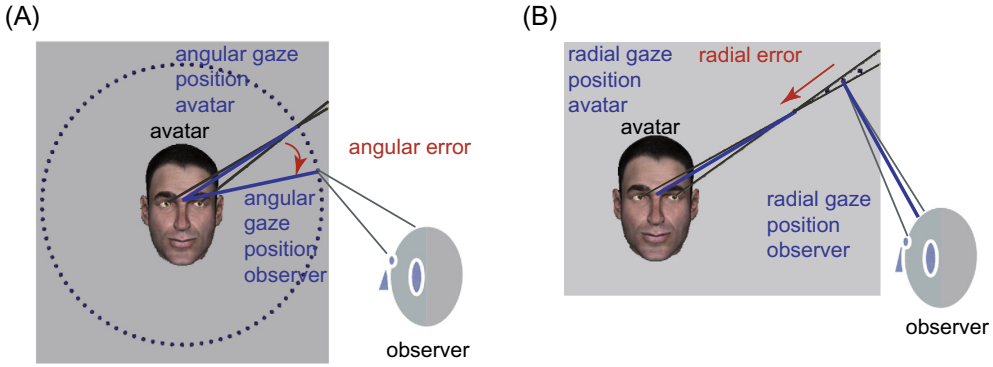

Fig. 1. Cartoons explaining the definitions of the angular (A) and the radial response errors (B) characterizing gaze following. The human observer is facing a human avatar at 80 or 100 cm distance with virtual target objects placed in a fronto-parallel plane midway between the two. The task of the observer is to indicate at which target the avatar is looking at.

difference later in the discussion. The subjects were instructed to shift a mouse cursor, (a stereoscopically presented ring (diameter = 1°) moving in the same depth plane as the target objects), on top of the target object, to which they perceived the avatar was looking at. The subjects were provided a temporal window of 20 s to deliver their response. Experiments 1, 2, 3 and 5 varied in terms of the spatial arrangement of target objects, which will be described in the methods section dedicated to the respective experiments. The sequence of events in a given trial corresponded to the one given for experiment 4 on the natural sender.

Prior to carrying out experiments 3 and 5, we subjected participants to a control experiment, which assessed whether the subjects actually perceived the stimuli stereoscopically. Moreover, we performed this control experiment retrospectively on 6 out of the 11 subjects who had participated in experiment 1 and 2, who were still available. In this control experiment subjects were presented two stimuli located left and right of straight ahead. Out of these two stimuli (spheres of 0.6° diameter each), the one on the right side was always kept in the same depth plane midway between observer and sender, whereas the left one could be either located at the same depth or 50, 150, or 250 cm before or behind this reference stimulus. The subjects were instructed to move the mouse cursor to the stimulus they perceived in the foreground. We carried out a probit analysis of the subjects' decisions (McKee, Klein, & Teller, 1985), in order to estimate the point at which they perceived both targets in the same depth plane. The probit analysis fits a sigmoidal function to the psychometric data, using a least-square algorithm. The parameters characterizing the probit function are its turning point and the slope at that point. The turning point of the probit function defines the point that yields equally frequent "left" and "right" decisions as the two targets are seen in the same plane. Those subjects whose responses yielded no significant fit ($p < 0.05$) were assigned to the "no-stereo-vision" group and the others to the "normal stereo-vision group".

2.4. Data analysis

In the experiments 1–5 we calculated the distribution of the observers' decision on the avatar's angular gaze position. We further determined both the response error and the absolute response error. The response error was given by the difference between the observers' estimate on which angular target position the avatar looked at ($°target_{(observer)}$) and the avatar's actual angular gaze direction ($°target_{(avatar)}$) (see Fig. 1A for a graphical explanation). We converted the resulting angular response error – and its absolute value respectively – into a visual response error. This was calculated by determining the Euclidian distance between the veridical and perceived target position on the circle. The specific details regarding data analysis will be detailed in the dedicated part for every experiment.

3. Experiment 1 – Angular gaze following

3.1. Participants

Eleven adults (age from 24 to 32 years, mean age of 26.7 years; 5 females and 6 males) participated in the first experiment as observers.

3.2. Experimental design

Experiment 1 (see Fig. 2A) probed gaze following prompted by targets showing a circular arrangement ("angular gaze following"). The human avatar shifted eye gaze to targets on a circle of 15° radius visual angle midway between himself and the observer. Ninety target spheres were equally distributed along the circle with an angular target spacing of 4° (from 0° to 356° in steps of 4°). Each of the 11 subjects completed one experimental session consisting of 90 trials, in which the order of the targets singled out by the avatar was pseudo-randomized. Before the start of the session all subjects reported that they perceived the target objects roughly midway between them and the human avatar. Targets were perceived in a depth plane 50 cm, respectively away from both the avatar and the human observer.

We also tested 6 subjects retrospectively, who had participated in experiment 1, in the aforementioned depth perception experiment. Based on the results, 3 subjects fell into the group with normal stereo-vision and the other 3 in the group with insufficient stereovision, a difference that was considered in the analysis of data from experiment 1.

3.3. Data analysis (experiments 1 and 2)

We describe the analysis of the data from experiments 1 and 2 jointly since they were very similar. In experiment 1 and 2 we calculated the distribution of the observers' decision on the avatar's angular gaze position. We further determined both the response error and the absolute response error. The response error was given by the difference between the observers' estimate on which angular target position the avatar looked at ($°target_{(observer)}$) and the avatar's actual angular gaze direction ($°target_{(avatar)}$) (see Fig. 1A for a graphical explanation). We converted the resulting angular response error – and its absolute value respectively – into

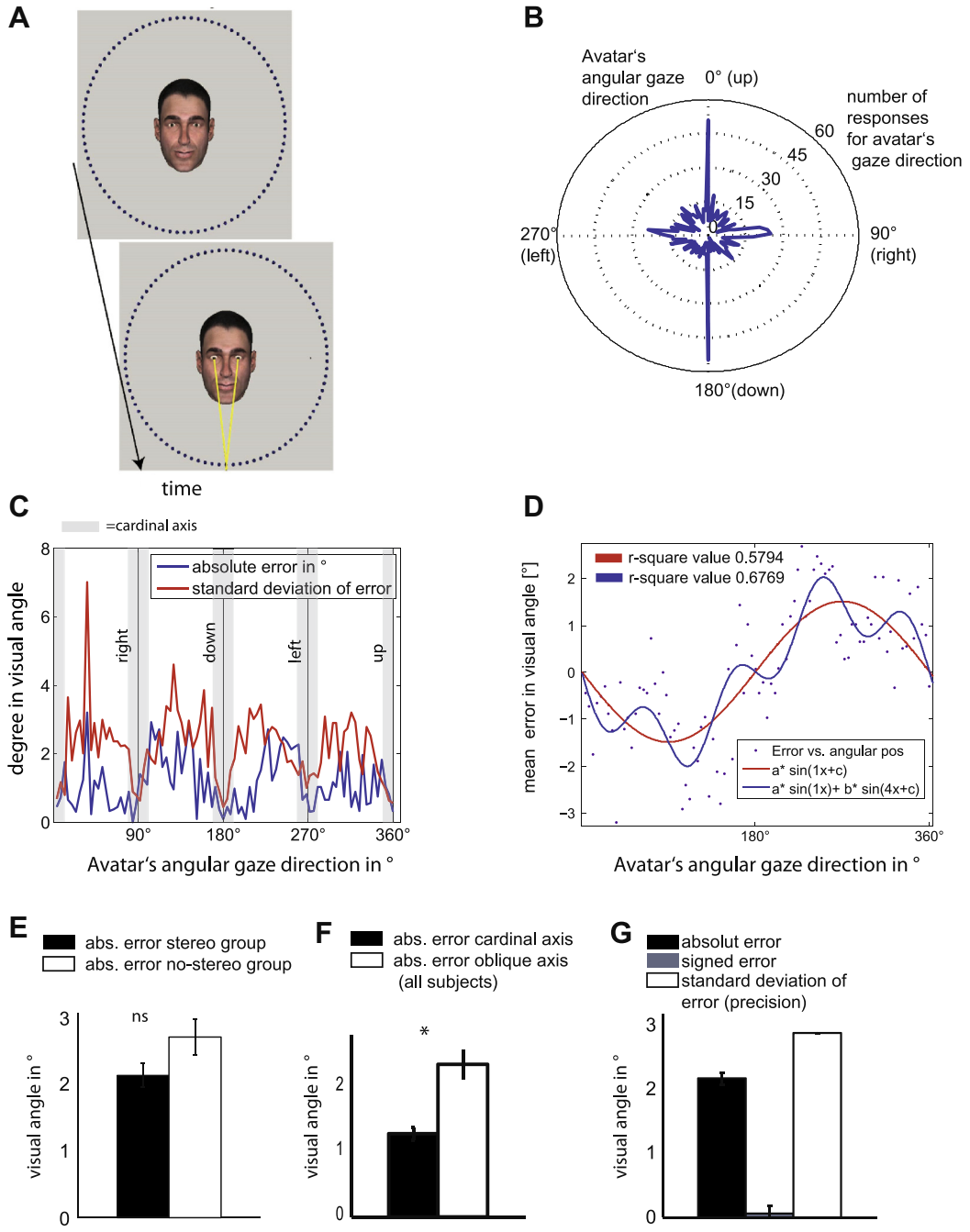

Fig. 2. Experiment 1. Panels B–G show data pooled over 11 subjects. (A) Design of experiment 1: One out of 90 target objects placed on a circle with a radius of 15° visual angle placed between the avatar and the observer are singled out by the avatar's eye gaze. The yellow lines (not displayed during the experiment) indicate the eye directions of the avatar fixating one of the targets, defining his gaze direction and the target hit by his gaze. (B) Distribution of target choices as a polarplot (angle in deg. visual angle, frequency of observations along the radius ($N_{(\text{num of observations})} = 990$ $n_{(\text{subjects})} = 11$). (C) Mean absolute angular response error in degree visual angle as function of direction (blue line) plus standard deviation (red line), ($N_{(\text{num of observations})} = 990$ $n_{(\text{subjects})} = 11$). (D) Fit of angular error with two harmonic functions (see inset for details) in an attempt to capture both the cardinal axis effect and the upward bias. Fitting with the first harmonic frequency only yielded a fit with an r-square value of 0.5794. Adding the fourth harmonic frequency yielded a better fit (r-square value 0.6769). (E) Mean averages of the absolute angular error and its standard error of the mean for the subset of the subjects ($n = 6$), which we could test on stereo-vision in order to assign them to the stereo-group and no-stereo group (rank-sum-test: n.s. $p = 0.1771$). (F) Mean average absolute response error plus standard error of the mean plotted separately for different axes (oblique and cardinal). The comparison revealed a significant influence of the axis: mean absolute angular error was 1.29° for all cardinal axes pooled vs. 2.49° for the pooled oblique axis (rank-sum-test: $p = 0.00044$). (G) Mean averages of the signed and absolute angular error and its standard error of the mean, and the standard deviations of the absolute response errors for all subjects. $N_{(\text{num of observations})} = 990$.

a visual response error. This was calculated by determining the Euclidian distance between the veridical and perceived target position on the circle. The resulting distribution of errors along the circle were fitted with a sine wave function a1 * sin(b1 * x + c) (using the CF toolbox in MATLAB 7.5.0). We fixed the value of b1 to 1 to test for the existence of an upward/downward bias of the responses, which appears as one sine-wave period over the circle. Furthermore we fitted the same data with a sum of two sine waves (a1 * sin(b1 * x) + a2 * sin(b2 * x + c)) in order to test for the existence of a bias aligned with the cardinal axes (setting b1 to 1 and b2 to 4) (Bock, Dicke, & Thier, 2008). In order to determine, whether there was a significant difference between the response

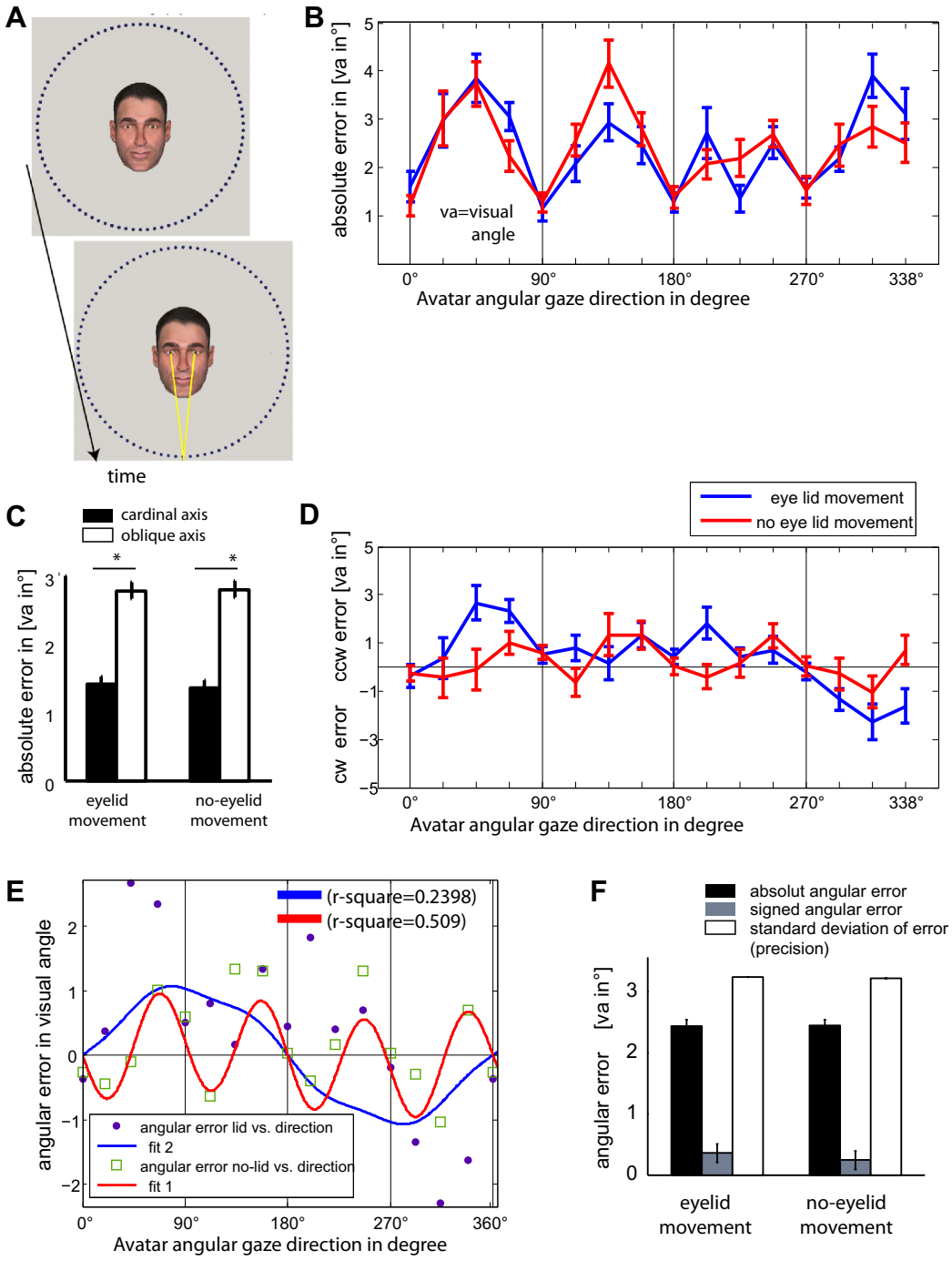

Fig. 3. Experiment 2. Panels B–F show data pooled over 5 subjects. (A) In experiment 2, the geometry of experimental setup was like in experiment 1, however we included conditions in which the avatar's eyelids moved with the eyes. (B) Plot of mean absolute angular error and standard deviation as function of gaze direction for experimental blocks with (blue) and without (red) eyelid movements (data for 5 subjects and per subject 90 trials for with- and without eyelid movement condition each), accompanying the gaze displacement. (C) Bar chart of mean absolute error plus standard deviation for the 2 experimental conditions eyelid movement vs. no eyelid movement) based on pooled data for the cardinal and the oblique axes respectively. (D) Plot of the signed angular error plus standard error for the with- (blue) and the without (red) accompanying eyelid movement conditions. (E) Plots of mean signed error as function of gaze direction for the with- and the without accompanying eyelid movement conditions, fitted with harmonic functions. The former was fitted with a sine function (r-square = 0.2398), the latter with the sum of this sine wave and its 4th harmonic frequency (r-square = 0.509). Refer to the text for explanation. (F) Bar chart showing the mean global absolute angular response error and standard deviation for the eyelid and no eyelid movement condition pooled across subjects and directions.

error on the cardinal axis and the ones on the other positions on the circle we defined sectors of 24° angle width, centered on the cardinal axis and compared the response errors between the cardinal sectors and the other sectors (oblique axis pools) by forming two pools, a cardinal sector pool and a non-cardinal sectors pool and subjected the resulting averaged values for a subject to a statistical analysis.

In experiment 2, in which we wanted to assess the influence of eyelid movements, we divided the circular target positions into 16 sectors, 4 of them centered on the cardinal axes (horizontal and vertical axis) and 12 on the 4 oblique axes. The sectors comprised between 5 and 6 response targets. We calculated the mean absolute response error in degree visual angle for every sector and for the two experimental conditions separately (presence and absence

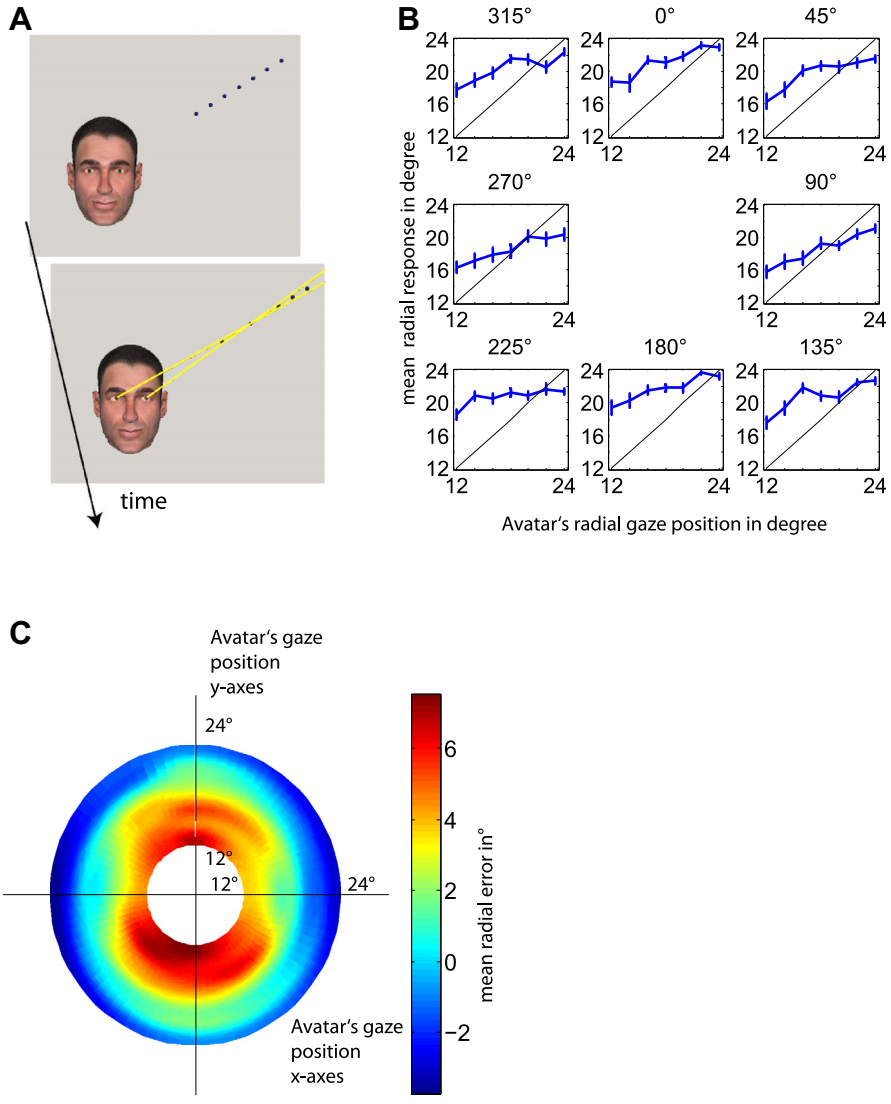

Fig. 4. Experiment 3. Panels B and C show data pooled over the 8 subjects with normal stereo vision. (A) Design of experiment 3: One out of seven targets (12–24° eccentricity in steps of 2°) placed on a rod of different orientations (from 0° to 315° in steps of 45°) in the plane midway between the two agents is singled out by the avatar's eye gaze. (B) Mean and standard error of the mean of subjects' target choices as function of eccentricity for different spoke orientations (avatar's gaze direction) ($n = 8$ subjects and each subject completing 116 trials). (C) Pseudo-3D plot of mean radial error as function of spoke orientation and eccentricity for different spoke orientations (gaze directions). The mean radial error is color-coded.

of eyelid movement, Fig. 3A blue and red line respectively). The absolute error measures the response error in visual angle with respect to the correct eye gaze position and does not discriminate between clock- and counter-clockwise angular errors.

3.4. Results

In experiment 1, subjects performed an angular gaze following task. In this experiment they had to indicate at which target object arranged on a circle the avatar was looking at. Subjects chose targets on the horizontal and vertical axis (cardinal axes) considerably more frequently than the other targets (see Fig. 2B), though the avatar looked at every target with equal frequency. The subjects were more precise for eye gaze directions addressing the cardinal axes targets as indicated by the distribution of the absolute error (|response° – target°|) for different targets on the circle (see Fig. 2C, blue line). In line with this observation the standard deviation of the absolute response error was larger for the oblique and smallest for the cardinal axis (see Fig. 2C, red line). We wanted to address the question, whether there is a significant difference in response errors on the cardinal relative to the oblique axes as shown by previous work using a naturalistic setting (Bock, Dicke, & Thier, 2008). To this end, we calculated for each subject the mean absolute response error separately for the cardinal and oblique axes pools and then subjected the resulting values to a statistical analysis. We found that the absolute error for the cardinal axis pool was significantly smaller than the one for the oblique axis pool (mean absolute cardinal error = 1.28°, mean absolute oblique error = 2.49°), (see Fig. 2F), rank-sum test, $p = 0.00044$.

We next addressed the question if the stereoscopic depth cue, provided in our setup, had an influence on the angular gaze following. To this end, we distinguished subjects with normal stereo vision (stereo group) from those with no or insufficient stereo vision (no-stereo group) (see Methods). Actually, we observed no significant difference in the size of the pooled absolute angular error between the stereo- and the no-stereo-group ($error_{(stereo)} = 2.15°$ and $error_{(no\text{-}stereo)} = 2.72°$, $p = 0.1771$, rank-sum-test; Fig. 2E). Also taking into account the signed

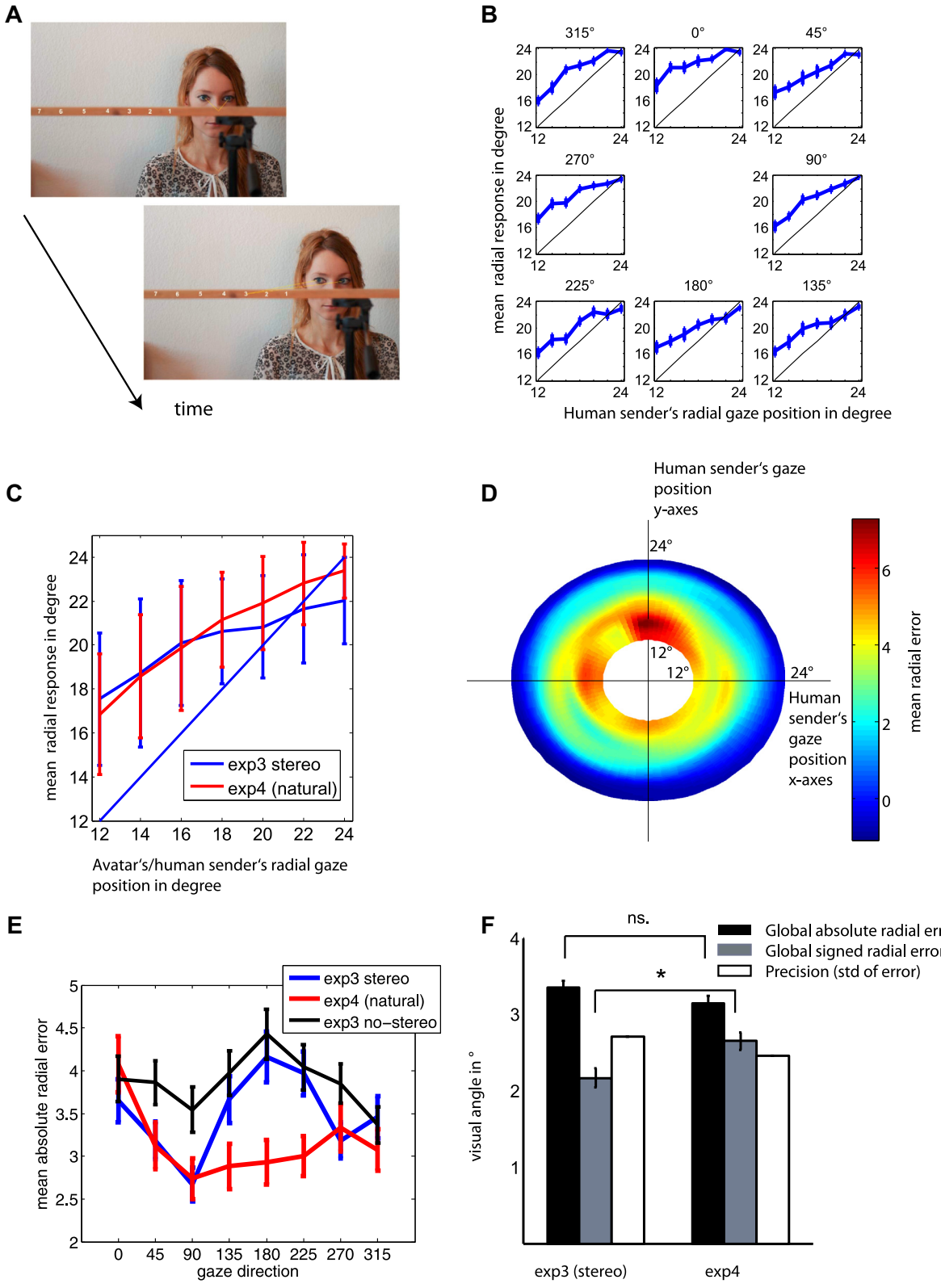

Fig. 5. Experiment 4. (A) Experiment 4 with the same geometry as experiment 3, but with human avatar being replaced by a human "sender" shifting eye gaze to real objects located on a wooden rod. (B) Mean target choices and standard error of the mean as function of eccentricity for different spoke orientations (human sender's gaze direction). Data for $n = 6$ subjects. (C) Mean target choices and standard deviation as function of the avatar's/human sender's gaze location, pooled across all spoke orientations in experiment 4 (red line; $n = 6$ subjects) and experiment $3_{(\text{stereo-group};\ n = 8\ \text{subjects})}$. The straight blue line denotes the unity line representing veridical responses. (D) Pseudo-3D plot of mean radial error as function of spoke orientation and eccentricity for different spoke orientations ($n = 6$ subjects). (E) Absolute radial error as function of the human sender's gaze direction for individual spoke orientations in experiment 4 (red line). In addition, for the purpose of comparison, separate plots are included distinguishing subjects with ($n = 8$) and without ($n = 11$) stereo-vision (blue and black line respectively). Error bars denote standard error of mean. (F) Mean global signed and absolute radial error and its standard deviation (precision) for experiments 4 and $3_{(\text{stereo-group})}$ (6 and 8 subjects respectively, rank-sum test).

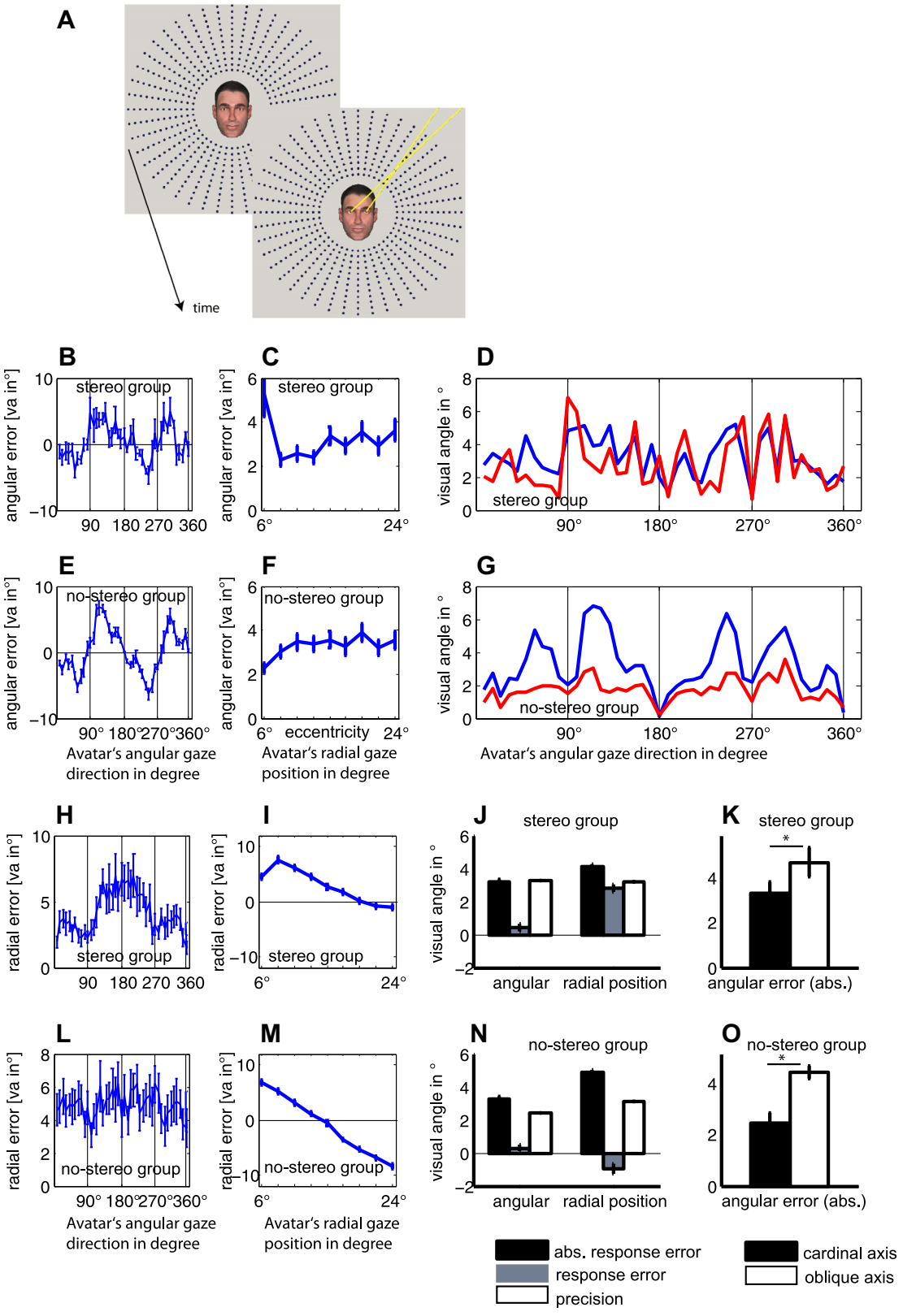

Fig. 6. Experiment 5. (A) Design of experiment 5: The avatar's eye gaze identifies one out of 528 targets distributed on 12 virtual concentric circles with radii ranging from 6° to 31°, with each circle containing 44 equally spaced targets. (B and E) Mean angular error as function of avatar's gaze direction for the stereo- ($n = 6$) and no-stereo-group ($n = 4$), respectively. (C and F) Mean angular error as function of avatar's gaze eccentricity for the stereo- and no-stereo-group respectively. (D and G) Absolute angular error and its standard deviation as function of avatar's gaze direction for the stereo- and no-stereo-group respectively. (H and L) Mean absolute radial error as function of avatar's gaze direction for the stereo- and no-stereo-group respectively. (I and M) Mean signed radial error as function of avatar's gaze eccentricity for the stereo- and no-stereo-group respectively. (J and N) Global values of the angular and radial error for the stereo- and no-stereo-group respectively. Result of statistical comparison is shown in Fig. 8B and in results section. (K and O) Angular error for targets on the cardinal and the diagonal axes for the stereo- and no-stereo-group, respectively.

response error yielded no significant difference between both groups ($p = 0.4403$; rank-sum-test). As the availability of normal stereo vision did not seem to make a difference, the following analytical steps were based on pooled data from both subgroups.

In order to reveal possible biases for the upward axis as reported by Bock, Dicke, and Thier (2008), we fitted the observed angular error from the target in degree visual angle with a sine function: angular error = a1 * sin (b1 * x) + c, fixing the frequency parameter b1 to 1. The red line (see Fig. 2D) indicates the shift of response errors in the upward direction; with a r-square value of 0.5794. The negative deviations for target positions 0–180° and positive deviations between 180° and 360° reflect a general upward shift of observers' estimates of where the avatar is gazing at. The observed smaller angular error around the cardinal axis could also be fitted with a sum of two sinus with the second one corresponding to the fourth harmonic of the first one, reflecting the observed bias of the cardinal axis: angular error = (a1 * -sin(b1 * x) + a2 * sin(b2 * $x + c$) with b1 and b2 fixed to 1 and 4, respectively) with a r-square value of 0.6769 (see Fig. 2D blue line) validating the bias of the responses for the cardinal axis ($p < 0.05$ for the reported fits). In summary, these results confirm the bias towards the cardinal axes as well as a general upward bias, shown previously in a naturalistic setup (Bock, Dicke, & Thier, 2008). The global (pooled across all conditions) absolute error, global error and the global precision (standard deviation of pooled errors) across all conditions and subjects were 2.17°, 0.05° and 2.89° visual angle respectively (see Fig. 2G). The precision observed in our experiment of 2.89° matches very well the values found by Bock et al. of 2.81° and 3.09° for the dynamic (subject sees the entire saccade from the beginning to the end) and static (subject does not see the saccade) condition respectively.

4. Experiment 2 – The influence of the eyelid information on angular gaze following

4.1. Methods

4.1.1. Setup generating 3D eye gaze avatars

In order to introduce additional eyelid movement cues in experiment 2, eyelid movements/positions, which corresponded to the avatar's eye gaze were generated in POSER ver. 7. We subdivided the avatar's eye gaze positions located on a circle in 9 horizontal sections stacked on each other, with the 5th one centered on the horizontal axis. Further, individual eyelid positions for each horizontal section were generated using POSER ver. 7. These considerations were based on the observation that purely horizontal eye gaze shifts do not change the eyelid position. These mesh images containing the different eyelid positions corresponding to the different gaze locations were used by the Virtual Gaze Studio software according to the eye gaze displacement of the avatar.

4.1.2. Participants

5 adults (age from 19 to 29 years, mean age of 25.4 years; 3 males) participated in the second experiment as observers.

4.1.3. Experimental design

Experiment 2 (see Fig. 3A), in which we used the same spatial arrangement of the targets as in experiment 1, addressed the question, whether the presence or absence of eyelid movements influenced the subject's ability to discriminate angular eye gaze position. To this end we tested 5 subjects' ability to discriminate the human avatar's eye gaze either lacking or exhibiting accompanying changes in eyelid position for changes in eye gaze orientation. The geometry of targets and the task instructions were the same as in experiment 1. Each subject completed two blocks of 90 trials each, in which we tested the two aforementioned conditions. The order of blocks was pseudo-randomized across subjects. The targets were perceived in a depth plane 50 cm away from both the avatar and the human observer. In every case, the avatar's eyelid position during the initial central fixation phase of a given trial was veridical, i.e. it was appropriate for the straight ahead orientation of the avatar's eyes (see Fig. 2A and 3A). Whereas the eyelids stayed in the same position throughout the remainder of the trial in the no-eyelid movement condition (see Fig. 2A), they changed their position according to the avatar's eye gaze displacement in the eyelid movement condition (Fig. 3A respectively).

4.2. Results

Using the same spatial configuration as in experiment 1, experiment 2 tested, if the presence or absence of eyelid movements influenced the accuracy of gaze following. First, we asked whether the accuracy of gaze following on cardinal and oblique sectors was influenced by the presence of eyelid movements. A two-way ANOVA with the factors availability of eyelid movement cues and sectors on the circle (cardinal and oblique axes) revealed a main effect of the sector ($p = 0.0106$), but no effect of the eyelid information availability and no significant interaction.

Calculating the mean absolute angular error for pooled cardinal and oblique sectors separately for each subject and running statistical analyses separately for the with and without eyelid movement conditions, revealed a similar mean angular error between conditions (with and without eyelid movement), with a significant difference between the cardinal and the oblique axes in both conditions (rank sum test, $p = 0.0079$ for no eyelid and $p = 0.0159$ for the eyelid movement present), see Fig. 3C.

Next, we asked whether there was an upward response bias as observed in experiment 1 and previous naturalistic experiments. The signed error exhibited a significant upward-bias in the eyelid movement condition only (see Fig. 3D and E, blue line) with large clockwise and counter-clockwise response errors close to the upward axis. Similar to experiment 1 we tried to fit the pattern of the signed error with sine functions of different frequencies. The best fit for the no eyelid movement condition was a sine wave with at the 4th harmonic frequency supporting a cardinal axes bias (r-square = 0.509) (see Fig. 3D, red line). This pattern is consistent with the first part of experiment 1 mentioned above. Overall, eyelid movement information did not improve accuracy of angular gaze following.

5. Experiment 3 – Radial gaze following with human avatar

5.1. Methods

5.1.1. Participants

19 adults (age from 19 to 49 years, mean age of 28.0 years; 11 males) participated in the third experiment.

5.1.2. Experimental design

Experiment 3 probed gaze following of targets showing a radial arrangement (''radial gaze following'', see Fig. 4A). In this experiment the human avatar looked at targets at varying radial eccentricities on a virtual spoke in a plane midway between the avatar and the observer. The virtual spoke was oriented in one of 8 possible directions from 0° to 315° in steps of 45°. On every spoke 7 spherical targets (diameter of 0.6°) were presented at eccentricities of 12–24° in steps of 2° (Fig. 4A). 19 subjects completed two sessions, with 56 trials each, in which the order of the targets was pseudo-randomized. The targets were perceived in a depth plane 40 cm away from both the avatar and the human observer.

Before carrying out experiment 3, we subjected participants to control experiment described in the general methods section, which assessed whether the subjects actually perceived the stimuli stereoscopically.

5.1.3. Data analysis (experiments 3 and 4)

In experiments 3 and 4 we determined the distribution of the observers' decision for the radial gaze position of the avatar and the human sender. Furthermore, we calculated the observers' mean responses for individual eccentricities (radial position) on different spokes. Additionally we computed the mean radial response error (error = $^{\circ}\text{target}_{(\text{avatar})} - {}^{\circ}\text{target}_{(\text{observer})}$) in degree visual angle. In order to test statistically the dependency of the absolute response error on the directions on the circle (0–315° in steps of 45°), we pooled for every subject in a group (e.g. stereo and non-stereo in exp 3) the absolute error across all eccentricities for each of the 8 directions and calculated the average for every direction.

5.2. Results

In experiment 3 the subjects performed a gaze following task, in which they had to indicate to which target object on a spoke the avatar was gazing at. Based on the aforementioned stereovision control experiment we could divide the subjects into a group with normal stereo-vision and one lacking stereo-vision. Both the stereo-vision group, consisting of 8 subjects, and the no-stereo-vision group with 11 subjects were subjected to the same analysis.

First we wanted to assess the influence of the eccentricity on radial gaze following. Therefore we calculated the mean observers' radial error relative to target position as a function of the position of the target singled out by the avatar for all spoke directions individually (0–315° in steps of 45°) for the stereo-group. We observed that subjects overestimated the avatar's eye gaze for smaller eccentricities and underestimated gaze for larger eccentricity independent of spoke orientation (see Fig. 4B and C). The transition of overestimation to underestimation occurred on average at 22°, where the gaze judgment revealed the least error (see Figs. 4B and 5C the blue line). Next we tested the influence of gaze direction on the discrimination performance. This was done by first averaging the absolute radial errors across all positions (12–24°) for every spoke direction (0–315°) as can be seen in Fig. 5E. The Kruskal–wallis test revealed that the distribution of the absolute response errors differed significantly from zero ($p = 0.0398$) for the stereo-group. For further analysis we considered the absolute error because averaging across the signed errors would falsely reveal a low average error in case subjects might have a bias for middle eccentric gaze positions. In such a case the positive and negative errors would cancel each other out. We observed an anisotropy of the amount of overestimation for the different directions, with the subjects' tendency to overestimate being the smallest around the horizontal axis (see Figs. 4C and 5E).

Next we wanted to assess whether perception of depth cues had an influence on the accuracy of radial gaze following. Therefore we compared the performance of the stereo-group with the no-stereo-group. For the non-stereo-group ($n = 11$) we observed throughout all directions higher absolute radial response errors compared to the group of subjects perceiving the stimuli stereoscopically ($n = 8$) and using depth cues to solve the task (see Fig. 5E black vs. blue line). The mean absolute radial error comprised 3.31° and 3.75° for the stereo and non-stereo-group respectively, and the comparison yielded a significant difference (rank-sum-test, $p = 0.021$ when determining for every subject the average absolute error for every direction and then comparing the averaged values across all subjects). As shown clearly, the usage of stereoscopic depth cues significantly improved the ability of the subjects to discriminate the radial gaze component.

6. Experiment 4 – Radial gaze following with human sender

6.1. Methods

6.1.1. Participants

The fourth experiment comprised six test subjects (age from 23 to 33 years, mean age of 27.6 years; 3 males) and one human actor, a blue-eyed female with no visual deficit, replacing the avatar.

6.1.2. Experimental design

In experiment 4 the trial started with a natural human sender, sitting face to face with the observer, looking straight ahead on the centre of a wooden rod containing the fixation target circle (diameter of 0.6°, positioned at the height of the sender's nasion). The human sender was a 27 year old caucasian female subject with no ocular imbalance. The human observers were asked to look at the same fixation target at which also the human sender looked. Upon a hint provided by a person standing behind the human observer, the human sender shifted eye gaze to one target midway between herself and the human observer (see Fig. 5A). The task of the human observer was to specify the number of the target on which he/she perceived the human sender to be looking at. The subjects were provided sufficient time (20 s) to indicate their response. They were allowed to break fixation of the central fixation spot and to follow the sender's gaze. A new trial was initiated once the response was delivered.

Experiment 4 like experiment 3 probed radial gaze following (Fig. 5A). However, as described earlier, the avatar was replaced by a human sender shifting eye gaze to objects placed on a wooden rod at different orientations. In this experiment the human sender provided eyelid and eyebrow movement cues unlike in experiment 3 using the avatar. Otherwise the spatial arrangement of the target objects was like in experiment 3 (see above). Six subjects with normal stereovision as assessed in our virtual reality setup completed two sessions each in which the order of the targets was pseudo-randomized.

6.2. Results

In order to rule out that the above mentioned results for radial gaze following were due to some peculiarities of our 3D-virtual reality setup using a human avatar we resorted to a control experiment (experiment 4) in which we replicated the above mentioned study (experiment 3), however replacing the avatar with a human sender (see Methods experiment 4). In this experiment the human sender also provided eyelid and eyebrow movement cues unlike in experiment 3 using the avatar. We wanted to assess whether we would confirm the observed pattern of over- and underestimation for smaller and larger eccentricities respectively on radial gaze following obtained with the human avatar in experiment 3. Fig. 5B shows the mean responses for individual spoke directions and eccentricities and reveals that subjects generally overestimated the radial gaze position like in experiment 3. The larger the eccentricity of the human sender's gaze the smaller the observer's deviation from the veridical radial gaze position was (Fig. 5B and C). As in experiment 3, averaging subjects' response deviation for individual directions, we observed an overestimation in the range of 2–4 degrees in visual angle (Fig. 5D). Unlike in experiment 3, in experiment 4 there was no significant effect of gaze direction (spoke direction) on the judgment of the radial position (Kruskal–wallis test, $p = 0.2774$), Fig. 5E red line.

We also wanted to assess the influence of eyelid movement information on the radial gaze following component, by comparing results of experiment 4, where the human sender provided eyelid movement information, with results of experiment 3, in which the avatar lacked these. The main difference between using an avatar with no eyelid movements (experiment 3) and a natural sender (experiment 4), can be seen in Fig. 5E. Subjects showed larger radial errors in experiment 3 for the direction 135–225°, where the avatar gazed downward compared to experiment 4, in which they followed the gaze of the human sender. Besides the downward direction, we did not observe any difference between experiments 3 and 4.

Despite these differences, both experiments using an avatar or a human sender revealed that subjects were generally biased to overestimate the radial eye gaze position (Fig. 5F gray bars). The comparison of the pooled absolute errors between experiment $3_{(\text{stereo-group})}$ and 4 revealed no significant difference (absolute error exp $3_{(\text{stereo})} = 3.31°$, absolute error exp $4 = 3.14°$, rank sum-test, $p = 0.34$). The standard deviation of all absolute errors was similar in both experiments, Fig. 5F white bars (2.71° and 2.45° for experiment $3_{(\text{stereo})}$ and 4 respectively).

7. Experiment 5 – Combined angular and radial gaze following

7.1. Methods

7.1.1. Participants

Ten subjects (aged from 24 to 49 years, mean age 29.4 years, 5 males) participated in the last experiment (5) as observers judging the eye gaze of a human avatar.

7.1.2. Experimental design

Experiment 5 tested gaze following abilities in two dimensions at the same time (Fig. 6A) by having the avatar shifting eye gaze to targets on a virtual sphere, whose radial (eccentricity) and angular (direction) location could be modified pseudo-randomly from one trial to another. These targets were located in a depth plane midway between avatar and subjects, 50 cm respectively away from both of them. This experiment comprised 528 targets, which the avatar could shift eye gaze to. Target position was defined by 12 radial positions at 44 different angular positions. Radial target position varied from 6° eccentricity to 31° with steps of 2.27°. Angular target position varied from 0° orientation to 351.82° with steps of 8.18°. Ten subjects completed one session each, in which the order of the targets was pseudo-randomized. As mentioned for experiment 3, we additionally tested these subjects' stereo-vision ability in our setup in a separate control experiment. Based on their depth perception ability we divided these 10 subjects into two groups: those with normal stereo-vision and those with insufficient stereo-vision group, in order to test the influence of stereo cues and the ability to use them for gaze following. Based on this independent criterion we assigned 6 subjects to the stereo- and 4 to the no-stereo-group.

In experiment 5, we considered only targets on the first 9 rings for the analysis and excluded targets on the 3 outermost circles as some subjects perceived them at a depth that differed from the one of the targets closer to straight ahead. This can be explained by the fact that we used a tangential screen, whereas the stereoscopic projection OpenGL deploys assumes a spherical screen. This change of depth plane became perceptually noticeable only for eccentricities exceeding 26°, i.e. in those eccentric parts of the visual field in which the 3 outermost target rings were located, which is why they were excluded from the analysis. In other words, the targets considered were located on circles ranging from 6° to 24.18°. In the other experiments 1, 2, 3, target eccentricities were limited to $\leqslant 24°$ right from the outset, and in no case did we obtain reports of deviant depth planes. Because of this rigorous eccentricity criterion eccentricity associated deviations of perceived depth cannot influence the results obtained.

7.1.3. Data analysis

In experiment 5 we calculated both the mean angular and the mean radial response error ($\text{error} = {}^\circ\text{target}_{(\text{avatar})} - {}^\circ\text{target}_{(\text{observer})}$) for all directions (from 0° orientation to 351.82° with steps of 8.18°, see Fig. 6A for illustration). Moreover, we determined the mean angular and radial response error for all radial target positions (from 6° to 24.18° eccentricity). In order to assess whether the size of angular or radial error was dependent on the cardinal or oblique axis we calculated for every subject the average absolute response error for the cardinal and oblique axis and ran statistical tests with these average values (horizontal and vertical axis determining the cardinal axis and the oblique axis on the left and right hemifield). In order to describe the cardinal-axis bias and the upward bias we fitted the mean angular response error over all eccentricities ($\text{error} = {}^\circ\text{target}_{(\text{avatar})} - {}^\circ\text{target}_{(\text{observer})}$) of the observers' estimate where the avatar looked at with a sine wave function as described above. We fitted mean angular errors with the function $a1 * \sin(b1 * x + c1))$ with $b1 = 2$ or, alternatively, with $a1 * \sin(b1 * x + c1) + a2 * \sin(b2 * x + c2)$ with b1 and b2 fixed to 2 (twice the fundamental frequency) and 4 (4th harmonic), in order to test for an upward/downward (vertical-axis) bias and cardinal-axis bias respectively.

We also calculated the 2D error vector, taking into account both the angular and radial error, for every target, averaging across all subjects. Finally, pooled measures of accuracy were computed and averaged across all target directions and eccentricities. They consisted of the pooled absolute error (across all subjects) the pooled signed error (=the response error with the sign of the error preserved vs. absolute, in which only the amount is considered) and the pooled precision value (=standard deviation of pooled absolute errors). The precision value mentioned above measures the variability of the responses and hence like the response error a small value indicates little variation of the responses and therefore high accuracy. These values were used to compare accuracy for different conditions in all five experiments. These global values enabled us to directly compare our findings with previous studies, which assessed gaze following.

7.2. Results

We conducted a fifth experiment in which the avatar shifted eye gaze from the center to one out of 528 objects arranged on 12 concentric circles, with radii of 6–31° in steps of 2.27°. Unlike the previous experiments, in which we either measured the angular component of gaze following (experiments 1 and 2) or the radial component (experiments 3 and 4), here we considered both the radial and the angular coordinates characterizing the observer's gaze position acquired in response to gaze of the avatar. In order to compare angular and radial gaze following of the stereo and no-stereo group and test the influence of stereoscopic cues we examined the response error for the angular and radial components over all subjects for the stereo- and the non-stereo-group separately. As can be seen in Figs. 6B and E and 7A and C we observed for both groups a characteristic pattern of clock and counter-clockwise error of the avatar's angular gaze position (clockwise and counter-clockwise error indicated by a positive and negative angular error respectively) centered on the cardinal axes. The pattern of clock- and counter-clockwise angular error with zero-crossing passing the cardinal axes reveals a main bias for the vertical axes (see Figs. 6B and E and 8A). When we analyzed the absolute angular error, we observed for both groups a larger

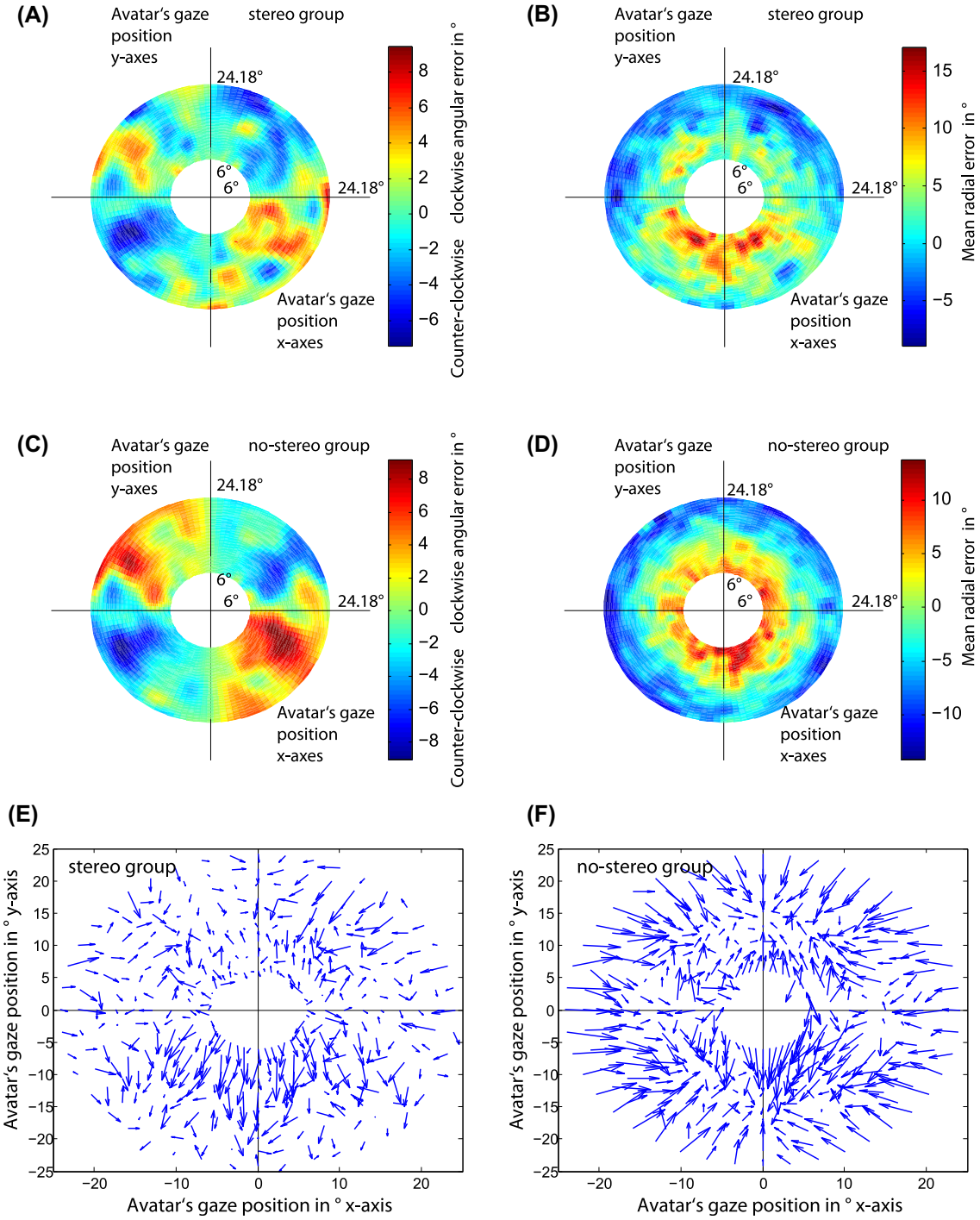

Fig. 7. Experiment 5. (A and C) Mean angular error distribution as function of avatar's gaze direction pooled over all subjects in the stereo-group ($n = 6$) and the no-stereo-group ($n = 4$) respectively. Warm and cold colors denote clock- and counter-clockwise deviations respectively of observers' choices ("error") from the actual avatar's gaze position. (B and D) Mean radial error distribution as function of avatar's gaze direction pooled over all subjects in the stereo-group ($n = 6$) and the no-stereo-group ($n = 4$) respectively. Warm and cold colors denote over- and underestimation of the avatar's radial gaze position respectively. (E and F) 2D-error vector field displaying the mean errors (combined radial and angular) for the stereo- and no-stereo-group, respectively. Each beginning of an arrow indicates the actual avatar's gaze position. The tip of the arrow shows the mean endpoint across all subjects.

angular response error for the oblique directions compared to the cardinal axes (Fig. 6D and G). In both groups the absolute angular error was significantly smaller for the cardinal axis compared to the oblique axis (angular error cardinal-axis$_{(stereo)} = 3.33°$ oblique-axis$_{(stereo)} = 4.73°$, angular error cardinal- axis $_{(non\text{-}stereo)} = 2.48°$ oblique-axis$_{(non\text{-}stereo)} = 4.47°$, Fig. 6K and O, $p = 0.0404$ and $p = 0.0019$ for stereo and no-stereo-group, rank-sum-test).

Next we wanted to test how larger eccentricities affect the angular error. We obtained larger angular error for the larger eccentricities (Fig. 6C and F). Furthermore, fitting the angular error with a sum of two sine waves having the second and fourth harmonic frequency, revealed the best fit with a r-square value of 0.63 and 0.82 for the stereo- and the no-stereo-group respectively. These results indicate that the both the vertical and the cardinal axes exerted a strong influence on the subjects (see Fig. 8A). In

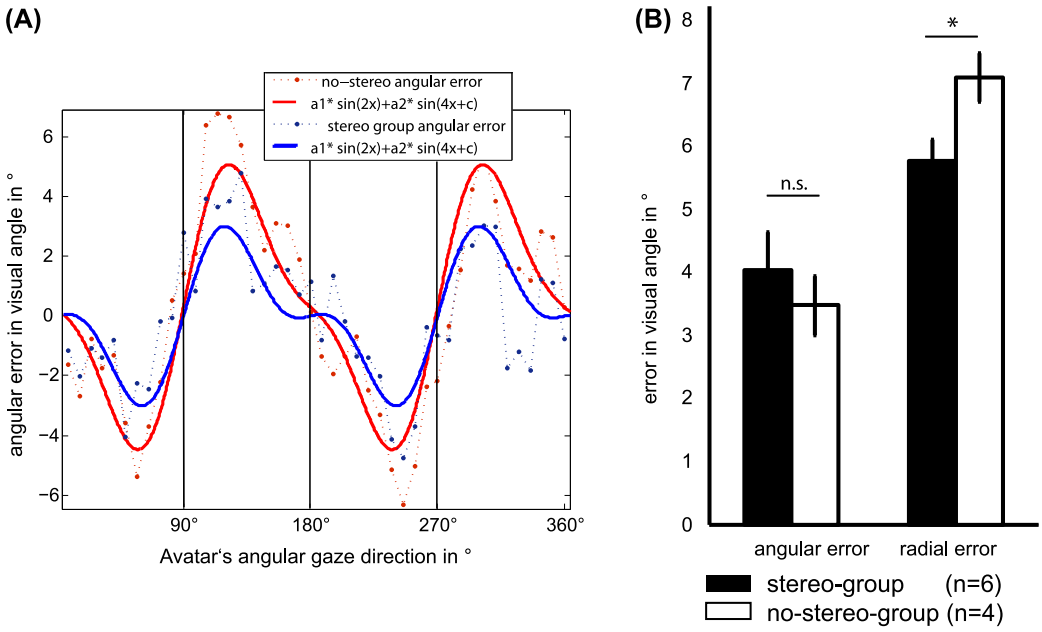

Fig. 8. Experiment 5. (A) Fit of mean angular error of stereo- (red) and no-stereo-group (blue) with combinations of sine functions (see inset for details). The dots connected by interrupted lines indicate the mean errors, the solid lines the fitted functions. (B) Comparison of the mean and standard errors of the mean absolute angular and radial errors for the stereo- and no-stereo-group.

order to test whether we would observe the same pattern of radial error, dependent on eccentricity, we calculated the response error for the different target locations across all subjects and uniformly observed overestimation for smaller radial gaze positions and smaller overestimation up to underestimation for the larger radial positions (see Figs. 6I and M and 7B and D, here shown for different gaze directions) in both groups. The radial error declined with larger eccentricities in the stereo-group, (see Fig. 6I) very similar to the results of experiment 3 and 4, in which we had assessed this component in isolation. Furthermore, the absolute radial error across all conditions was smaller in the stereo-group compared to those who could not use stereo cues to judge the radial position (rank sum-test, $p = 0.0039$, see Fig. 8B). In the non-stereo-group the radial response was compressed to a small range. In addition in the stereo group the radial error was significantly smaller in the horizontal axes compared to the oblique and vertical axis (see Figs. 6H and 7B and E). We also assessed whether we would observe a similar increase in the radial error in the downward direction as in experiment 3 due to the lack of eyelid movements in experiment 5. The larger radial error around the downward direction confirms that the lack of eyelid movements being prominent in these direction has a substantial influence on the radial gaze discrimination ability in humans. The radial gaze position overestimation was the largest in both groups, when the avatar's gaze was directed downward, (see Fig. 7B, D, E, and F). Next we visualized the combined angular and radial error across all subjects (see Fig. 7E, F; here every arrow starts with the avatar's gaze position and the direction and the length of the arrow denotes the subjects' combined 2D response error). Across all gaze directions we observed larger radial errors for subjects who could not use the depth cues provided in our setup (Fig. 7E and F). Furthermore as mentioned above the radial error was largest for the downward gaze direction. With respect to the angular gaze component we observed a global angular error (absolute) of ~3.7° visual angle which was larger than in the first experiment (~2.3°), with a similar standard deviation of ~3° compared to 2.88° in the first experiment. Similarly, we observed a larger radial error$_{(absolute)}$ in this last experiment compared to both the experiments 3 and 4, in which we had assessed the radial component in isolation. Experiment 5 confirms that lack of stereo cues hampers radial gaze following and further suggests that detecting both the angular and radial gaze position simultaneously is more demanding than one of these components in isolation.

8. General discussion

Despite the ecological importance of gaze following, little is known about the underlying neuronal processes, which allow us to extract gaze direction from the geometric features of the eye and head of a conspecific. In order to understand the neuronal mechanisms underlying this ability, a careful description of the capacity and the limitations of gaze following at the behavioral level is needed. Previous studies of gaze following relied on naturalistic settings (i.e. the observer following the gaze of a real human person either performing in front of the observer or shown in a naturalistic portrait, shifting her/his gaze to locations and/or objects). This setting has the disadvantage of allowing only very limited control of potentially relevant visual features guiding gaze following, such as the contrast of iris and sclera, the shape of the eyelids or any other geometrical features and – in the case of photographs – lacking depth. Hence, in order to get full control of potentially relevant features we decided to study gaze following of human observers guided by the gaze of a human avatar ('sender') seen stereoscopically. To this end we established a stereoscopic 3D virtual reality setup, in which we tested human subjects' abilities to detect at which target a human avatar was looking at. The target was virtually placed in a fronto-parallel plane midway between the human subject and the human avatar. The target to be chosen from a set of objects was singled out based on the eye gaze of the avatar looking at that target object. In a series of 5 experiments we addressed the subjects' abilities to use the avatar's eye gaze direction to identify targets presented in varying locations on a circle ("angular gaze following"), at varying eccentricities on designated radial axes ("radial gaze following") or in locations corresponding to the nodes of a grid in the fronto-parallel plane, defined by combinations of angular and radial coordinates. We also tested the influence of eyelid

movements and the quality of observers' stereo-vision on the precision of gaze following. Independent of the particular condition, following the gaze of the avatar showed all the features of the following of the gaze of a natural person, namely, in general a substantial degree of precision as well as a consistent pattern of deviations from the target, characteristic for the particular condition prevailing. Poor stereo vision affected performance only in the case of targets not confined to positions on the circle. Only gaze following guided by targets at larger downward eccentricities exhibited a differential effect of the presence or absence of accompanying movements of the avatar's eyelids and eyebrows.

8.1. Methodological considerations

8.1.1. Interpupillary distance

The interpupillary distance of human observers is known to vary (Gordon et al., 1989). In our subject group the interpupillary distance amounted to a mean value of 64.12 mm and a standard deviation of 2.4 mm. On the other hand, when generating the 3D scenery we assumed a fixed interpupillary distance of 60 mm throughout in all our experiments. Could this slight difference have influenced the results? Subjects with a larger interpupillary distance than the 60 mm used in the experiments 1, 2, 3 and 5 would perceive the target plane not right in the middle between them and the avatar, but a little bit further away towards the avatar. In case the subject's interpupillary distance would be 6 mm larger than the one used by our visualization software, the target plane would be perceived to be located approximately 5 cm further away from the subject towards the avatar. The resulting change of the position of the target object position would amount to about 10% visual angle. Hence, a target located at 18° eccentricity would be mislocalized by about 1.8°, i.e. less than the smallest distance (2° in experiment 3) between targets. In 10 out of the 12 subjects tested, the deviation of the actual interpupillary distance was smaller than the one assumed in the preceding scenario. In other words, we may be confident that the assumption of a standard interpupillary distance is a reasonable simplification not leading to relevant errors.

8.1.2. Reports on target choice

A further principal methodological concern may be that our subjects deployed different *types of responses* in order to indicate their target choices. Whereas in the case of a natural sender they provided a verbal report, in all experiments with avatars they indicated their chosen target by moving a computer mouse-controlled virtual ring over the chosen target. The fact that we did not find a difference between comparable experiments deploying avatars and natural human senders (exp 3 vs. 4; exp 1 vs. experiments of Bock, Dicke, & Thier, 2008) speaks against an influence of response mode.

8.1.3. Duration of gaze shifts

A final concern may be the highly artificial *temporal structure* of the gaze shifts in the case of the avatar. The avatar shifted gaze from one position to the other from one frame to the next, whereas a human sender generates a saccade, which develops within a finite duration based on several frames. Again, the fact that the performance of observers in the two experiments comparing avatars and natural human senders was fully comparable suggests that this difference may actually be irrelevant. This does not mean that the visual system may not use information on the temporal structure of the change in eye position. Actually, Bock, Dicke, and Thier (2008) showed that the observer is able to detect gaze orientation better if observers are allowed to observe a saccadic shift from one position to the other. However, our findings suggest that this benefit may be largely based on a comparison of snap shots of the eyes at the starting and the end positions rather than information on the detailed kinematic structure of the intervening saccade.

8.1.4. Acuity

The smallest gaze shift of the avatar's eyes took place in experiment 3 in which subjects had to retrieve target location on radial axes. The distance of neighboring targets on the horizon amounted to 1.32 arcmin visual angle, corresponding to 28 pixels for targets presented along the horizontal axis. Assuming that our subjects had a visual acuity of 1, they should have been able to discriminate spatial offsets of 1 arcmin, i.e. offsets smaller than the smallest one offered in any of the experiments.

8.1.5. Proportion of subjects exhibiting stereo-vision in our setup compared to total subject participating

A share of 40–50% subjects exhibiting impaired stereo vision as found in our experiments greatly exceeds the prevalence of stereo-blindness reported in the literature (Coutant & Westheimer, 1993; Richards, 1970). Our classification was based on a statistical fit of the subjects' response distribution with a probit function on a limited number of data points. This is a very rigorous demand that will inevitably lead to an overestimation of the number of stereo blind subjects as those exhibiting more variable responses may fail to yield statistically significant fits although their stereo vision may in principle be good. Yet, this conservative approach has the advantage of allowing us to be sure that those who passed the criterion have indeed perfect stereo vision. In other words, while our approach does not allow us to draw conclusions on the true prevalence of impaired stereo-vision, the risk of falsely assuming stereo vision in our experiments is probably negligible.

8.1.6. Cardinal axis bias and symmetry

Our subjects exhibited a remarkable response bias for the cardinal axis. The reasons remain speculative. For instance, it is well known that both the visual system and the oculomotor system exhibit features that differentiate cardinal axis properties from non-cardinal axis features. For instance, cardinal axis saccades are known to be faster than oblique saccades (King, Lisberger, & Fuchs, 1986) and visual orientation discrimination is better for the cardinal axes (Appelle, 1972; Howard, 1982), in both cases reflecting features of the neuronal basis of these functions. Yet, it is not obvious why these features should lead to a response bias and other, unidentified factors may be causative.

Besides the observed cardinal axis preference our avatar's face exhibited perfect left–right symmetry. Natural faces exhibit a certain asymmetry in their visual features, whose relevance for gaze following has to our best knowledge never been addressed. This difference between the avatar's face and the human face might explain why we observed such a sharp response peak for the vertical axis in the experiments 1 and 2 using the avatar: one could speculate that the perfect left–right symmetry of the avatar's face influences the angular responses of the subjects and bias them towards the vertical axis. Although in experiment 1 we observed the cardinal axes bias and similar amount of global angular response precision like in the naturalistic experiment (Bock, Dicke, & Thier, 2008), the vertical axis exerted a larger influence than the horizontal axis.

In future experiments we should critically challenge this hypothesis by generating an avatar face with less symmetrical facial features and test whether we observe a diminished influence of the vertical symmetry axes.

8.2. Comparison with previous studies using naturalistic settings

Gaze following in our experiments was characterized by systematic response errors whose size and direction depended on target eccentricity in the case of radial gaze following, and the angular relationship relative to the cardinal axes in the case of angular gaze following. Moreover, gaze following to a given target position

showed a certain degree of response variability which was on the order of ~3°. An early study by Cline (1967) looked at gaze following along the two cardinal axes, guided by the reflected mirror image of a human sender. The author found mainly overestimation of target location, largely decreasing with increasing eccentricities. However, in the downward direction the overestimation increased with eccentricity. Similarly, we found an overestimation that depended on eccentricity with the largest overestimation for small eccentricities and little overestimation – or even reversal to underestimation for the largest eccentricities. As the size and direction of this systematic mislocalization was independent of the sender being an avatar or a natural person or (exps. 3 vs. 4), we may be confident that this finding does not reflect insufficiencies of the avatar stimulus. The reason for the consistent tendency to overestimate target locations close to the fixation spot remains unclear. In any case the decrease of this bias with target eccentricity may reflect the tendency of observers to generate hypometric saccades when trying to approach targets at large eccentricities.

Also another study, by Anstis, Mayhew, and Morley (1969), found an overestimation of target positions, yet only for targets on the horizontal axis and actually considerably larger with positive dependence on eccentricity (Anstis, Mayhew, & Morley, 1969; e.g. ~10° overestimation for 20° target eccentricity). Interestingly the overestimation in the study by Anstis and coworkers was diminished when the human sender was replaced by an artificial eye, exposing artificial sclera without coverage by eyelids, arguably facilitating the extraction of eye gaze direction based on information provided by the pupil. This result supports the notion that the eyelids matter when trying to extract gaze direction. It is important to note that in the experiment by Anstis and coworkers the removal of eyelids seemed to be beneficial. As noted earlier we found conversely that the addition of eyelids, taking over natural positions on the eyes when making gaze shifts, helped. The reason for this discrepancy might be related to their usage of an artificial eye.

9. Conclusions

In summary, we have shown that human gaze following guided by the eyes of a 3D avatar shows most of the features that characterize gaze following induced by the eyes of a natural person. This conclusion suggests that avatars, offering almost infinite possibilities to manipulate potentially relevant features may be useful when trying to more precisely delineate the cues exploited by observers. An example of the usefulness of this approach is our demonstration of the importance of eyelid movements accompanying the gaze shifts. Other interesting factors that may be addressed in future experiments should be changes in facial expressions, the role of shading, facial symmetry and last but not least interactions with head position.

Acknowledgments

This work was supported by a Grant from the Deutsche Forschungsgemeinschaft (TH 425/12-1). Further, we would like to acknowledge the work of Friedemann Bunjes supporting us technically and further Patrick Benz, Constantin Hübner and Malte Schomers developing the human avatar program.

References

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31*, 657–674.
Ando, S. (2004). Perception of gaze direction based on luminance ratio. *Perception, 33*, 1173–1184.
Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television "portrait" is looking. *American Journal of Psychology, 82*, 474–489.
Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation Oblique effect in man and animals. *Psychological Bulletin, 78*(4), 266–278.
Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press, p. 32.
Benz, P., Hübner, C. (2008). *Virtualisierung von Experimentumgebungen zur perzeptionellen Blickrichtungsbestimmung: Bestimmung des Referenzsystems der visuellen Verarbeitung von Blicken* (M.A. thesis). Faculty of Information and Cognitive Sciences, Tuebingen University.
Bock, S. W., Dicke, P., & Thier, P. (2008). How precise is gaze following in humans? *Vision Research, 48*, 946–957.
Cline, M. G. (1967). The perception of where a person is looking. *American Journal of Psychology, 80*, 41–50.
Coutant, B. E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthalmic and Physiological Optics, 13*(1), 3–7.
Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*, 519–526.
Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*, 581–604.
Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*, 694–724.
Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76*, 386–394.
Gordon, C. C., Bradtmiller, B., Clauser, C. E., Churchill, T., McConville, J. T., Tebbetts, I., et al. (1989). *1987–1988 Anthropometric survey of U.S. army personnel: methods and summary statistics*. TR-89-044. Natick MA: U.S. Army Natick Research, Development and Engineering Center.
Howard, I. P. (1982). *Human visual orientation*. Wiley.
King, W. M., Lisberger, S. G., & Fuchs, A. F. (1986). Oblique saccadic eye movements of primates. *Journal of Neurophysiology, 56*(3), 769–784.
Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387*, 767–768.
Langton, S. R., Watt, R. J., & Bruce, I. I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4*, 50–59.
McKee, S. P., Klein, S. A., & Teller, D. Y. (1985). Statistical properties of forced-choice psychometric functions: implications of probit analysis. *Perception & Psychophysics, 37*(4), 286–298.
Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition, 77*, B1–B14.
Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research, 10*(4), 380–388.
Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266.
Sinha, P. (2000). Last but not least. Here's looking at you, kid. *Perception, 29*, 1005–1008.
Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology, 131*, 451–469.
Todorovic, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research, 46*, 3549–3562.
Wollaston (1824). On the apparent direction of eye in a portrait. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 114*, 247–256.
World Medical Association (1964). Declaration of Helsinki. *British Medical Journal, 818*.