



# Impact of Iris Size and Eyelids Coupling on the Estimation of the Gaze Direction of a Robotic Talking Head by Human Viewers

François Foerster, Gérard Bailly, Frédéric Elisei

## ► To cite this version:

François Foerster, Gérard Bailly, Frédéric Elisei. Impact of Iris Size and Eyelids Coupling on the Estimation of the Gaze Direction of a Robotic Talking Head by Human Viewers. Humanoids 2015 - IEEE-RAS 15th International Conference on Humanoid Robots, Nov 2015, Séoul, South Korea. pp.148-153. hal-01228887

**HAL Id: hal-01228887**

**<https://hal.science/hal-01228887>**

Submitted on 14 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Impact of Iris Size and Eyelids Coupling on the Estimation of the Gaze Direction of a Robotic Talking Head by Human Viewers<sup>\*</sup>

François Foerster<sup>†</sup>, Gérard Bailly<sup>‡</sup>, Frédéric Elisei<sup>§</sup>

**Abstract**— Primates - and in particular humans - are very sensitive to the eye direction of congeners. Estimation of gaze of others is one of the basic skills for estimating goals, intentions and desires of social agents, whether they are humans or avatars. When building robots, one should not only supply them with gaze trackers but also check for the readability of their own gaze by human partners. We conducted experiments that demonstrate the strong impact of the iris size and the position of the eyelids of an iCub humanoid robot on gaze reading performance by human observers. We comment on the importance of assessing the robot's ability of displaying its intentions via clearly legible and readable gestures.

## I. INTRODUCTION

One important challenge of social robots is to ease interaction by complying with the rules that have been built by thousands of years of evolution of human societies. One aspect of this challenge is thus to provide robots with skills for analyzing human behaviors and generating relevant multimodal behaviors given their limited sensory-motor and cognitive resources. We claim here that the care given to the generation of multimodal behaviors should be accompanied with the assessment of its final rendering by the robot.

Human-robot face-to-face interaction (HRI) is in fact a coupled dynamical system where sensory-motor loops regulating the context-sensitive and situation-aware behaviors of both partners are mutually influenced. Robots should pick-up relevant characteristics of multimodal signals from the human partners and the surrounding environment in order to motivate actions. This challenge is eased by the so-called helpfulness of humans. In a series of books [1], [2] Tomasello and his colleagues notably propose that the phylogenetic specificity of humankind rests in its species-specific adaptation for sociability. The account offered by Tomasello contrasts human cooperation and altruism with nonhuman primate competition, and proposes that human altruism leads to shared intentionality (the ability to share attention to a third object and, more generally, to share beliefs and intentions). Tomasello [3] further proposed the cooperative eye hypothesis (CEH) that we will develop below. The CEH suggests that the eye's visible characteristics evolved to ease gaze following.

Conversely in HRI, humans expect social robots to offer back cooperation, altruism and share goals and plans. Such cooperative behavior will also be favored by the robots' gaze readability. This readability is both a control and design issue: the robotic eyes should be controlled and move in an appropriate and predictive way but should also be designed so that the eye's visible characteristics are similar to those that humans have developed for the sake of social interaction.

This paper describes one experiment demonstrating that eye geometry and texture matters. Taking liberties with these characteristics is risky and may lead to prejudicial misestimations of the gaze direction by human partners. We found that the relative proportion of pupil/iris/sclera and the coordination of eyelids and eye elevation in the built robotic eyes strongly impacts gaze estimation by human observers.

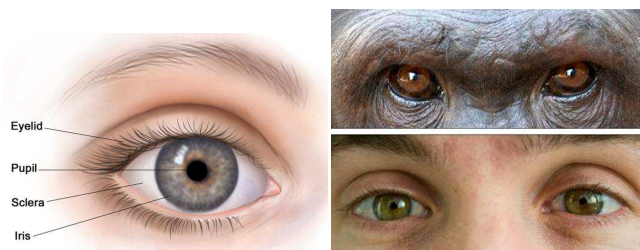


Figure 1: Left, human eye appearance (reprinted from cancer.gov); right: comparing the dark sclera of a chimpanzee with the strong contrast of the human eye.

## II. STATE OF THE ART

We compare here the structure and appearance of the human eye. We then review the structure and appearance of the eyes of emblematic social robots.

### A. Human eyes

The human eye ball is almost spherical in structure. According to [4], the size of the eyeball at birth averages 16.5 mm in diameter (front to back measurement). In adults, the diameter is 24.2 mm and that maximum eye size is reached when a person is 7-8 years old. The template aging effect [5] is mainly due to changes in iris texture.

The eye is not shaped like a perfect sphere; rather it is a fused two-piece unit. The smaller frontal unit, more curved, called the cornea is linked to the larger unit called the sclera. The corneal segment is typically about 8 mm in radius. The cornea and sclera are connected by a ring called the limbus. The iris – the color of the eye – and its black center, the pupil, are seen instead of the cornea due to the cornea's transparency.

<sup>\*</sup> Research supported by the ANR (Robotex ANR-10-EQPX-44-0 and SOMBRERO ANR-14-CE27-0014) and the UJF (EMSOC 2012084RECF991)

<sup>†</sup> François Foerster is a MSc student from Univ. of Grenoble-Alpes, currently with GIPSA-Lab, Univ. Grenoble-Alpes/CNRS, France (e-mail: francois.foerster@gipsa-lab.fr)

<sup>‡</sup> G. Bailly is a senior CNRS Research Director at GIPSA-Lab, Univ. of Grenoble-Alpes/CNRS, France (e-mail: gerard.bailly@gipsa-lab.fr)

<sup>§</sup> F. Elisei is a CNRS Research Engineer at GIPSA-Lab, Univ. of Grenoble-Alpes/CNRS, France (e-mail: frederic.elisei@gipsa-lab.fr).

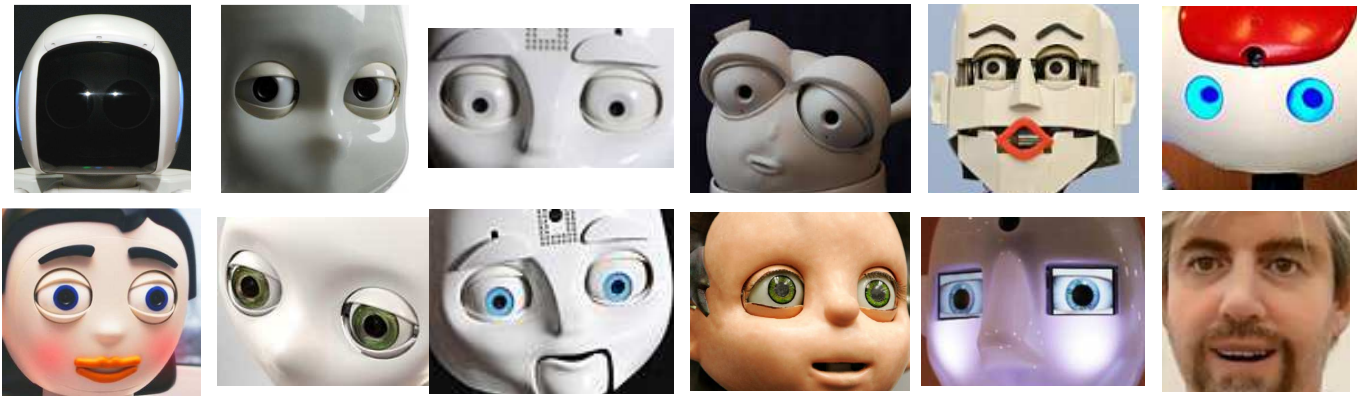


Figure 2: A gallery of robotic eyes. From left to right, top to bottom: hidden gaze of Asimo; original iCub, Nexi, Reeti and Kobian with no iris, Nao with no sclera; Flobby, Meka, Nexi and Zeno with sclera, iris and pupil; the Robothespian eye graphics vs. the hyper-realistic eyes of Geminoid DK. The panel of anatomies and pupil/iris/sclera proportions is very diverse.

Kobayashi et al [6] compared the width/height ratio of the eye outline (WHR) and an index of exposed sclera size in the eye outline (SSI) for different species: humans have the highest WHR ( $1.95 \pm 0.34$ ) and SSI ( $2.9 \pm 0.91$ ). The cooperative eye hypothesis [3][6][7] proposes the human eye's distinctive visible characteristics evolved to make it easier for confederates to follow another's gaze while communicating or working together on joined tasks. In fact, unlike other primates, human beings have eyes with an enhanced color contrast between the white sclera, the colored iris, and the black pupil (see Figure 1). This is due to a lack of pigment in the sclera. Kobayashi et al [6] hypothesize that the human eye, with a large sclera surrounding the iris and a great ability of eyeball movement, has provided a favorable ground for a drastic change from “the gaze-camouflaging eyes” of primates into “the gaze-signaling eyes” through a small change in scleral coloration. This gaze signal is supposed to be one of the main determinants of social interaction: most theory of mind (ToM) models [8][9] in fact postulate a specific neural mechanism (eye direction detector - EDD) in human brains specialized to detect others' eye directions. This EDD is one of the basic modules of the intentionality detector (ID) that enable us to attribute and compute others' beliefs, desires and intentions (see notably the BDI model introduced by Leslie et al [9]).

### B. Robotic eyes

Most humanoid robots do have eyes (see Figure 2). These eyes are not always active: cameras are often embedded in the pupils but sometimes deported in the nose or the forehead to get a more stable field of vision (cf. Nexi, Robothespian...). Most robots freely depart from the human pupil/iris/sclera textures and structure: while the Nao opts for a colored sclera with no iris, the Reeti has only pupils and scleras. Note that most robots have double eyelids, often contributing equally to the blinks and eye closure.

### C. Perception of the robot's gaze by human observers

A great amount of work has been dedicated to robot visual perception and multimodal scene analysis. Very high performing algorithms are today available to estimate gaze directions of human partners by robots. These estimations may strongly contribute to the computation of conversational regimes [10] and serve as input cues for generating their own gaze patterns for handling mutual attention and turn-taking

[11][12]. On the reverse, there are surprisingly few works assessing the perception of robots' gaze by human observers (see [13] for the Nao). The complexity of robot-centric research (communication, dexterous manipulation, navigation...) has somehow put the “readability” of the robot's actions into the background. To our knowledge, no work has been similarly reported on the audiovisual intelligibility of speech – i.e. the benefit of facial animation to the comprehension of the spoken message – uttered by robots. Remarkable experiments have nevertheless been conducted first by Delaunay et al.[14] with lamp avatar: they compared the estimations of the gaze direction of four different types of facial interface: a real human face, a human face displayed on a flat-screen monitor, an animated face projected on a semi-sphere and an animated face projected on a 3D mask. They show that robot faces having a human-like physiognomy and, surprisingly, video projected on a flat screen perform equally well. Instead of interposing a transparent grid between the participants and the faces, Al Moubayed et al [15][16] replicated part of these experiments with a chessboard using a lamp avatar called Furhat. They compared the estimation by human observers of the gaze of Furhat, its virtual model and the video of its performance both displayed on screen. They showed that Furhat is the only display escaping from the Mona Lisa effect [17] and delivering very accurate gaze direction, independently of the observer's viewing angle. If these experiments evidence the strong impact of 3D display, they do not provide insights into the possible contributing factors in eye geometry and texture that possibly modulate the resulting precision. Onuki et al [18] compared the impression given by mechanical eyes and a lamp avatar: they concluded that eyes with a round outline shape and a large iris were most suitable for precision and subjective agreement.

This paper describes a series of experiments demonstrating that eye geometry and eyelid positions matter to gaze estimation. We found that the relative proportion of eyelid/pupil/iris/sclera in the built robotic eyes impacts gaze estimation by human observers.

## III. NINA: A TALKING & GAZING ICUB

The original iCub robot is a 1 meter humanoid robot test bed for research into human cognition and artificial



intelligence. It was designed by the RobotCub consortium of several European universities and built by the Italian Institute of Technology [20]. In its final version, the robot has 53 actuated degrees of freedom. The eyes embed stereo cameras in a swivel mounting with similar degrees-of-freedom as human eyes (azimuth, elevation and vergence).

IIT and GIPSA-Lab recently collaborated to redesign the iCub head in order to enhance its communicative skills [19]. A new eye tilt mechanism – together with the suppression of lower eyelids – allowed obtaining the space for a Visaton BF32 loudspeaker and five additional motors (see Figure 3) for articulating the jaw and the lips. We chose to cover this mechatronic articulated head with a stretch fabric. The tissue covers the whole head with only four holes for the neck, the eyes, the top back of the skull (for ventilation) and the lips. The lips’ thickness was obtained by sewing two layers of stretchy fabric whose corners are screwed on the rods driven by the four lip motors. This new design also encloses two high-quality ear microphones plugged into the auricles.

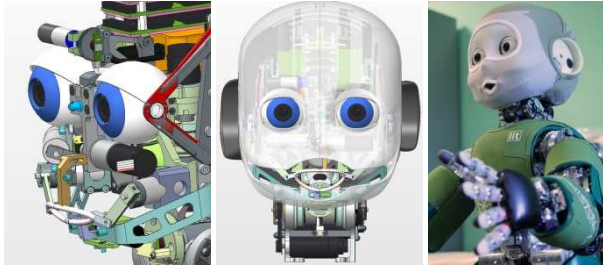


Figure 3. Details of the iCub talking head. Removable eye caps are highlighted in blue. Left: four motors are shaping the lips. The lower lip is carried by the jaw. Center: the skull and jaw consist of two rigid covers. Right: Nina’s final appearance.

#### A. Gazing

It is well-known that the subjective estimation of gaze direction and expression depends not only on the relative position of the pupil with the sclera but also on multiple other cues such as head direction [20], shape of the eyelids and palpebral borders [21] as well as eyebrows [22]. During our first teleoperation experiments, we were however quite surprised that the remote users complained about inaccurate eye fixations despite the very accurate control (parallax of the cameras, calibration of the eye tracker, centering of the visual field<sup>\*\*</sup>, etc.) of our perception/action loop. We thus decided to conduct a series of experiments on the estimation of gaze direction.

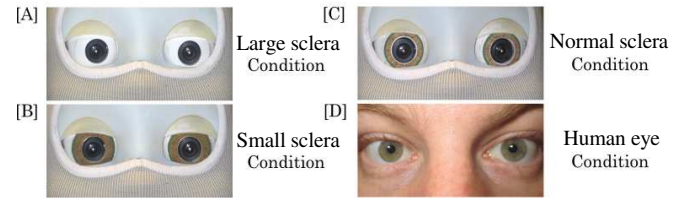
#### IV. EXPERIMENT: ESTIMATING POINTS OF INTEREST OF HUMAN VS. ROBOTIC EYES

The experimental environment extends the one used by Imai et al [23] for the evaluation of monoscopic vs. stereoscopic display systems.

<sup>\*\*</sup> Saccades of the robot’s eyes are triggered by estimating the fixation point from the binocular view and centering this point via a multilinear mapping that associates stereo fixation points to eye positions (azimuth, elevation and vergence).

We designed three different pairs of the removable ocular caps surrounding the eye cameras of our robot (see Figure 3) to evaluate the impact of the eyes features in a gaze-reading task (see Figure 4 for the appearance of the eyes and Table 1 for dimensions). The first pair is white (see Figure 4A), so the eyes are similar to the original iCub eyes. The second pair is fully textured thanks to brown iris-like decals (see Figure 4B). Finally, the decals of third pair get a smaller textured surface so that the final proportion of textured iris and white sclera in the eyes is similar to human eyes (see Figure 4C). To compare these robotic caps with natural eyes, we added a human-to-human interaction HHI condition where a human experimenter plays the same role as the robot (see Figure 4D).

##### Eyes features Hypothesis



##### Eyelids adjustment Hypothesis

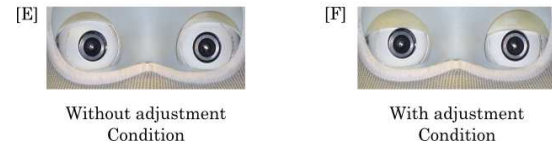


Figure 4. Robot’s and human’s eyes and robot’s eyelids used for the experiment, accordingly to our hypotheses. [A] white caps; [B] fully iris-like textured caps; [C] caps half white and half iris-like textured; [D] human eyes; [E] robot’s eye gaze without adjustment of eyelid position; [F] robot’s eye gaze with eyelid position adjusted. [E] and [F] are here featured with white caps.

Table 1. Dimensions (mm) and features of the three pairs of exchangeable caps surrounding the eye-embedded cameras

Feature	Large sclera	Normal sclera	Small sclera
Inner diameter of the caps	18		
Outer diameter of the caps	30		
Outer diameter of the iris	18	23	30
Percentage of iris in the caps	0	50	100

Concerning eyelid movements, we implement two different behaviors in our task. In one hand, upper eyelids were always static and opened at the same position, independently of gaze orientation (see Figure 4E for condition [A] without eyelid adjustments). In the other condition, the upper eyelid positions were constantly adjusted to be as low as possible without occluding the cameras’ field of view (see Figure 4F). Note that the 3D printer produces sand-colored eyelids, close to the iCub skin color.

#### A. Experimental setup

The task for the participants was to estimate which tile of a chessboard the agent (robot or human) was looking at (see Figure 5). The chessboard was placed on a table between the participant and the agent. In order to control distances of perception, the distance between front feet of chair’s participants and chessboard’s center was 35 cm. Distance between eyes of the agent and center of the chessboard was

set to 70 cm. The height of the table was 75 cm. The chessboard was a 61.5 cm x 61.5 cm plastic tablecloth with a small cross at the center of each tile. It was composed of 10 x 10 black and white tiles and each tile measured 5.1 cm by sides. The  $x$  axis shows letters from A to J in columns and the  $y$  axis displayed numbers from 0 to 9 in rows. Instruction was to estimate which one of the tile the agent looked at. Crosses' tiles looked by the experimenter or the robot were randomly chosen among the 64 center's tiles (B to I and 1 to 8). Because we wanted to authorize overestimation and underestimation of gaze orientation, peripheral tiles (i.e. columns A & J and lines 0 & 9) were neither gazed by the robot nor by the human experimenter.

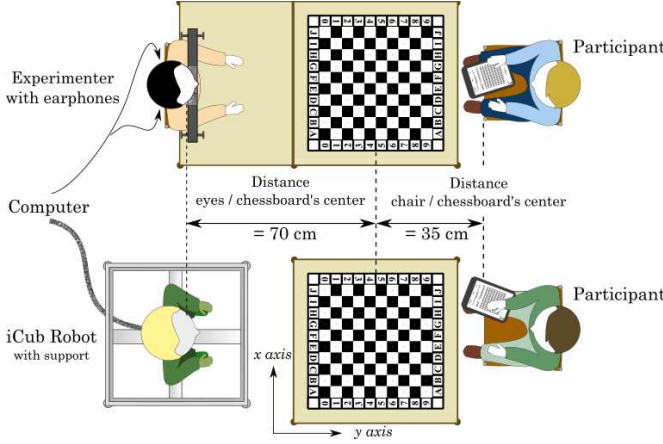


Figure 5. Top: the human-human interaction setup. Bottom: the human-robot interaction setup (see Figure 6 for front views)



Figure 6. On the left, the HHI setup with the chin strap. On the right, the HRI setup with the iCub robot.

**HRI condition.** The distance separating eye's robot and the ground was 125~cm. In these conditions, our iCub robot automatically chose and looked at each cross of tile. In order to validate the efficiency of the gaze program and the calibration, the experimenter verified on-line on a computer display that each tile was well pointed by the iCub's cameras. Furthermore, an off-line verification - by a laser pointer attached to the optics of the robot's cameras - was performed to be sure that mechanical calibration of gaze pointing was accurate. The robot's head position was fixed and consistent through trials and subjects. After each saccade performed by the robot (without an intermediate fixation towards the centre of the chessboard), subjects hear a tone and the trial's number to avoid response shifting.

Then, they have 10 seconds to report the estimated tiles on a paper notepad (see Figure 6).

**HHI condition.** Distance separating the human eyes and the ground was set to 115 cm. To control the possible impact of the head orientation on gaze reading, the experimenter placed his chin on a chinrest (see Figure 6). The tiles chosen randomly by the system were communicated to the experimenter by earphones. Just as in the HRI setup, the tone and the trial's number were heard by the participant who has also 10 seconds to report.

Participants were informed about the random order of the conditions (i.e., they will perform the task four times with all of our three different robot's caps and with eyes' experimenter in random order) and the fact that the targeted tiles are randomly chosen (i.e., 64 tiles among the 100 are randomly chosen by the program). They were not informed that peripheral tiles are never looked at. In addition, they were not aware of the eyelid position adjustments that were again randomly activated. Lightings in the room were always the same to control their possible impact on gaze perception.

## B. Subjects and dependent variables

In our study, 21 naive subjects (11 males and 10 females) took part in the experiment. The age of participants varied between 18 and 29 years (mean age=22.3 for males and 23.1 for females). Sex and age proportion were balanced. All of them had normal or corrected-to-normal vision and volunteer to carry out the experiment for free.

Each participant performed the task with the three couples of caps and the human eyes condition. The ordering of these four conditions were randomized for each subject. In the HRI condition, 32 targets were pointed with adjustments of the eyelid positions and 32 other targets without. The order of these targets was then totally randomized with the constraint of having all successive gazes at least two tiles apart. Because the eyes of the experimenter were not manipulated, we consider that the human eyes condition was performed with the adjustment of the eyelid positions.

In summary, we have two independent variables: eyes' characteristics (three couples of caps and human eyes) and eyelid dynamicity (eyelids with or without adjustment to the vertical position of the eyeball). Our experiment design followed a 4 (eyes characteristics: human eyes, *Small-sclera*, *Large-sclera* vs. *Normal-sclera* caps) x 2 (with vs. without adjustments of the eyelid position) factors in a within-subjects plan. Our dependent variable is the error distance between looked tiles and estimated tiles. We also evaluate separately the errors of azimuth and elevation angles: we also calculated separately eyes' misreading in  $x$  axis (i.e., column's error) and  $y$  axis (i.e., row's error), as in [23].

## C. Results

We applied a series of analysis of variance with the  $R$  program. The significance level for all tests was set to  $p < .05$ . Figure 7 displays error maps for our 7 gaze reading conditions. The height of the maps is proportional to the mean amplitude of the errors. We observe that (a) errors are mainly related to the distance between agent's eyes and target tiles, whatever the conditions and independently of the adjustment of the eyelid positions: largest errors were made

on far tiles from the agent's eyes and smallest errors concern close tiles (see also Figure 8); (b) this bias seems lower with adjustment of eyelids to the eyeball positions, particularly when the robot wore *Large-sclera* caps; (c) furthermore, a global effect of gaze misreading appears on azimuth peripheral tiles. This phenomenon is also observed for human eyes. The distribution of errors observed for human eyes is quite similar with those obtained with robot caps composed of *Small-sclera* and *Normal-sclera*, but appears different from those with *Large-sclera* where errors are mainly distributed at the right side of the agent. Finally, global errors of gaze-direction estimation are lower with the adjustment of eyelids. In the next section we confirm these first observations with statistical analyses of global errors.

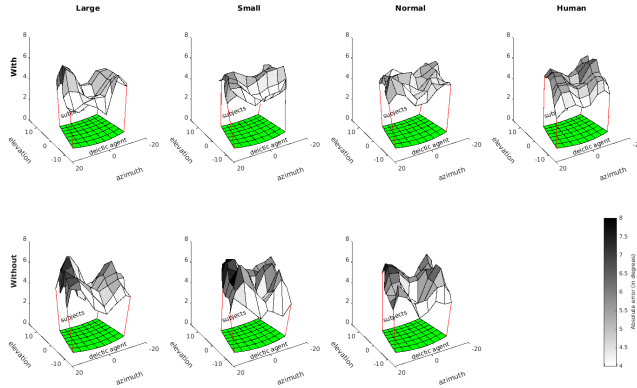


Figure 7. Maps of global error distribution for azimuth and elevation angles. These maps have been filtered by 2D-DCT (retaining 36 coefficients) to remove high frequencies. Gray levels indicate the amplitude of error related to each looked tile on the chessboard. Eye characteristics are displayed in column and adjustment of eyelids in line.

#### D. Stats

We performed an analysis of variance (AoV) of the global errors with the following predictors:

1. eye characteristics (4-ary)
2. eyelid adjustment (binary)
3. sex of the subject (2-ary)
4. elevation of the robotic right eye towards the target tile (continuous although 64-ary)
5. absolute azimuth of the right eye towards the target tile (continuous although 32-ary, considering left/right symmetry of the error functions)
6. side: (2-ary: left/right)

All predictors have a significant contribution to the variance of global error ( $p < e-6$ ). We obtain significant interactions between elevation and all other factors except side ( $p < e-3$ ). We also get interactions between absolute azimuth and eye characteristics elevation (see Figure 8) as well as. The only other significant interaction occurs between eyelid adjustment and eye characteristics with side ( $p < e-3$ ).

We performed AoV both on angles and tiles. We got the same results. The results for angular positions are given next. Post-hoc Tukey analyses revealed that, when eyelids are not adjusted, *Normal-sclera* caps produce significantly lower azimuth errors than other caps,  $p < 0.06$  (see Figure 8) while *Large-sclera* caps result in significantly better estimation of

elevation. Moreover, with adjustments, errors are significantly less sensitive to azimuth and elevation with *Normal-sclera* than other caps (see Figure 7). Concerning analyses of human eyes condition, tests indicated that global errors were significantly higher with human eyes compared to all caps in HRI,  $p < e-7$ . Note finally that females produced significantly better estimates than males (see Figure 8).

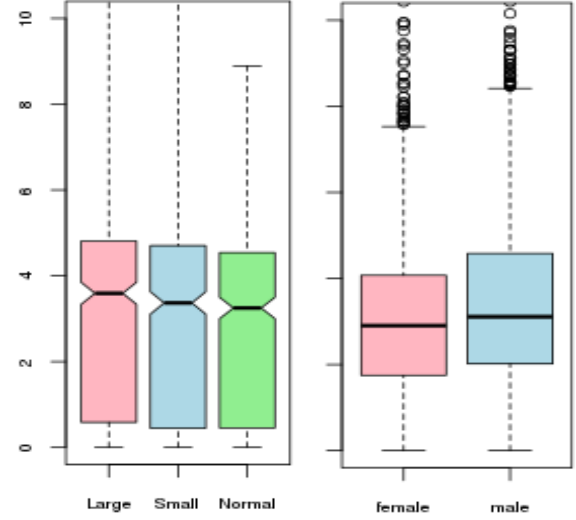


Figure 8. Left: Impact of sclera size on abs(azimuth error) with no eyelid adjustment; right: female performed significantly better than males.

#### V. DISCUSSION

Our analyses support the idea that the eyelid dynamicity improves the gaze readability. Global errors and elevation as well as azimuth errors are lower with adjustments of the eyelid positions. Our results are opposed to [24] who did not find any significant effect of the eyelid dynamicity on the accuracy of the gaze reading.

*Large-sclera* caps with the high contrast induced a more accurate estimation in the elevation dimension than the other caps. A possible explanation is that the composition of the caps with textured iris induced a weaker perceptual contrast compared to the eyes solely composed of a white sclera and black cameras. But surprisingly, these *Large-sclera* caps are not so efficient regarding the azimuth reading. The high perceptual contrast into the eyes is however not the only feature to examine. We would have to consider other features like the color of the skin and even eyelashes. For humans, the sclera can be discerned from the iris, eyelids and skin and it is in agreement with the “gaze-signaling” adaptive mode of the CEH [3]. As seen in Figure 4, the iCub exhibits a very small color difference between the sclera and the skin-mask. Further investigations will be crucial to know whether the skin color of the robot's head - and notably the external white hemline of the facial cover - impedes the gaze perception. The impact of side deserves comments: this could be due to lightening differences (i.e. the soft lighting from the semi-occluded windows at the left side of the agents in Figure 6).

Unexpectedly, the robotic caps induced lower gaze misreading errors than the human eyes in HHI. However, comparative results between HHI and HRI gaze reading tasks should be taken with caution since we had only one reference



gazer for our HHI experiment. Hence, all results found in the human eyes condition can be explained by the intrinsic characteristics of the experimenter's eyes morphology (e.g., exposed surface, eyebrows length, etc) as well as dynamic behavior. We need also to mention that our chin strap did not allow us to adjust the elevation angle of the head in front of the chessboard. Consequently, while the iCub's head was always turned toward the center of the chessboard, the experimenter's head was facing the participant. Several publications pointed out the influence of the direction of the head on the eyes perception [20], [25]. Consequently, during the task, the perception of the experimenter's eyes might be less efficient because of the head orientation.

Last limitation is about the robot eyes dimensions. Our iCub robot has exposed eyes with 30 mm in height and 35 mm in width: this is much bigger than the human eyes. Obviously, we can also assume that gaze perception depends on the size of the eyes (i.e., the bigger the eyes, the easier the gaze perception) and consequently reconsider Onuki et al results [18]. According to Kobayashi and Kohshima's analyses, humans have the highest mean width/height ratio of the eye outline (WHR close to 2.0 - see §II.A) while the iCub's eyes ratio is 1.17. Future research needs to dissociate the size of the eyes and the width/height ratio in order to take into account the influence of these features in gaze perception.

## VI. CONCLUSIONS

We assessed the cooperative eye skills of our talking head. We tried to demonstrate that interaction should be grounded on robust encoding and decoding modules. We have confirmed that gaze availability has a non negligible impact on communication efficiency. Several other factors should be examined such as the impact of head directions of both the robot and its reader as well of other aspects of eye morphology (eye shape and size, pupil size, iris color, etc) and other facial components (eyebrows, skin color, etc). A robotic platform offers valuable opportunities for studying factors affecting the eye reading capabilities of humans (and other species).

We thus plan to assess in the same way other communication skills (speech production, hand pointing...) before engaging these modules in autonomous HRI experiments.

## ACKNOWLEDGMENTS

Miquel Sauze conducted preliminary experiments on gaze estimation. Nina's head was designed in close collaboration with IIT and fabricated in Genova. We receive invaluable hardware & software support from IIT and want to especially thank Marco Ranzato for his availability. We warmly thank our subjects and the GIPSA-Lab mechatronics studio (in particular Julien Minet and Ghatfan Hasan) for keeping NINA alive.

## REFERENCES

- [1] M. Tomasello, *Origins of Human Communication*. Boston, MA: MIT Press, 2008.
- [2] M. Tomasello, *Why We Cooperate*. Cambridge, MA: MIT Press, 2009.
- [3] M. Tomasello, B. Hare, H. Lehmann, and J. Call, "Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis," *J. Hum. Evol.*, vol. 52, pp. 314–320, 2007.
- [4] D. Vaughan, T. Asbury, and P. Riordan-Eva, *General ophthalmology*. Stamford, Conn.: Appleton & Lange, 1999.
- [5] S. P. Fenker and K. W. Bowyer, "Analysis of template aging in iris biometrics," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 45–51.
- [6] H. Kobayashi and S. Kohshima, "Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye," *J. Hum. Evol.*, vol. 40, no. 5, pp. 419–435, 2001.
- [7] H. Kobayashi and S. Kohshima, "Unique morphology of the human eye," *Nature*, vol. 387, pp. 767–768, 1997.
- [8] S. Baron-Cohen, D. A. Baldwin, and M. Crowson, "Do children with autism use the speaker's direction of gaze strategy to crack the code of language?," *Child Dev.*, vol. 68, no. 1, pp. 48–57, 1997.
- [9] A. M. Leslie, "ToMM, ToBY, and Agency: Core architecture and domain specificity," in *Mapping the Mind: Domain specificity in cognition and culture*, L. A. Hirschfeld and S. A. Gelman, Eds. Cambridge: Cambridge University Press, 1994, pp. 119–148.
- [10] K. Otsuka, "Conversation Scene Analysis," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 127–131, 2011.
- [11] A. Mihoub, G. Bailly, and C. Wolf, "Modelling perception-action loops: comparing sequential models with frame-based classifiers," in *Human-Agent Interaction (HAI)*, Tsukuba, Japan, 2014, pp. 309–314.
- [12] D. Bohus and E. Horvitz, "Decisions about turns in multiparty conversation: From perception to action," in *ICMI*, pp. 153–160.
- [13] R. H. Cuijpers and D. van der Pol, "Region of eye contact of humanoid Nao robot is similar to that of a human," in *Social Robotics*, vol. 8239, G. Herrmann, M. J. Pearson, A. Lenz, P. Bremner, A. Spiers, and U. Leonards, Eds. Springer International Publishing, 2013, pp. 280–289.
- [14] F. Delaunay, J. Greeff, and T. Belpaeme, "A study of a retro-projected robotic face and its effectiveness for gaze reading by humans," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Osaka, Japan, 2010, pp. 39–44.
- [15] S. Al Moubayed, G. Skantze, and J. Beskow, "Lip-reading: Furhat audiovisual intelligibility of a back-projected animated face," *Intell. Virtual Agents - Lect. Notes Comput. Sci.*, vol. 7502, pp. 196–203, 2012.
- [16] S. Al Moubayed, J. Edlund, and J. Beskow, "Taming Mona Lisa: communicating gaze faithfully in 2D and 3D facial projections," *ACM Trans. Interact. Intell. Syst.*, vol. 1, no. 2, p. article 11 (25 pages), 2012.
- [17] R. Gregory, *Eye and Brain: The Psychology of Seeing*. Princeton, NJ: Princeton University Press, 1997.
- [18] T. Onuki, T. Ishinoda, E. Tusburaya, Y. Miyata, Y. Kobayashi, and Y. Kuno, "Designing robot eyes for communicating gaze," *Interact. Stud.*, vol. 14, no. 3, pp. 451–479, 2013.
- [19] A. Parmiggiani, M. Randazzo, M. Maggiali, F. Elisei, G. Bailly, and G. Metta, "An articulated talking face for the iCub," in *Humanoids*, Madrid, Spain, 2014.
- [20] D. Todorović, "Geometrical basis of perception of gaze direction," *Vision Res.*, vol. 46, no. 21, pp. 3549–3562, 2006.
- [21] A. Casari, F. Elisei, G. Bailly, and S. Raidt, "Contrôle du regard et des mouvements des paupières d'une tête parlante virtuelle," in *Workshop sur les Agents Conversationnels Animés*, Toulouse - France, 2006, pp. 1–9.
- [22] R. Watt, B. Craven, and S. Quinn, "A role for eyebrows in regulating the visibility of eye gaze direction," *Q. J. Exp. Psychol.*, vol. 60, no. 9, pp. 1169–1177, 2006.
- [23] T. Imai, D. Sekiguchi, M. Inami, N. Kawakami, and S. Tachi, "Measuring gaze direction perception capability of humans to design human-centered communication systems," *Presence*, vol. 15, no. 2, pp. 123–138, 2006.
- [24] S. Al Moubayed and G. Skantze, "Perception of gaze direction for situated interaction," in *Workshop on Eye Gaze in Intelligent Human-Machine Interaction (Gaze-In)*, Santa Monica, CA, 2012, p. Article 3 – 6 pages.
- [25] N. J. Emery, "The eyes have it: the neuroethology, function and evolution of social gaze," *Neurosci. Biobehav. Rev.*, vol. 24, no. 6, pp. 581–604, 2000.