Robot Gaze Is Different From Human Gaze: Evidence that robot gaze does not cue reflexive attention

Henny Admoni
Department of Computer Science
Yale University
New Haven, Connecticut USA
henny@cs.yale.edu

Brian Scassellati
Department of Computer Science
Yale University
New Haven, Connecticut USA
scaz@cs.yale.edu

ABSTRACT

Research has shown that, on a behavioral level, robot gaze can elicit social responses typically reserved for interactions with people. On the other hand, we have found that on a lower, reflexive level, robots cue attention differently than social symbols like human faces. This paper presents a popular psychophysics method for measuring attention that is novel to HRI. Using this attention cueing paradigm, and following previous psychology studies showing that faces and eyes are powerful social cues, we examined the attentional effects of robot gaze, human gaze, and socially neutral directional symbols. We discovered that while human faces (including line drawings of human faces) and arrows cued reflexive attention shifts, robot gaze did not. These results suggest a difference in low-level processing of robot versus human gaze, which stands in contrast to similarities found in high-level processing of robot and human gaze.

Keywords

gaze, psychophysics, human-robot interaction

1. INTRODUCTION

Evidence suggests that robot gaze can convey meaningful information to people. Robots are able to define conversational roles for human partners, such as addressee, bystander, or eavesdropper, entirely through visual attention cues [7]. People engaged in a guessing game are susceptible to subconscious attention cues from a robot that shifts its eyes briefly to the intended target, with a humanoid robot showing more success at directing peoples' attention than a robot with a more stylized appearance [8].

Although behavioral evidence shows that robots are perceived as social entities, we have found evidence that robots are treated differently from other social symbols on a lower, reflexive, perceptual level. Our experiments reveal that on a time scale of hundreds of milliseconds, robot gaze does not cue the same automatic attention shifts seen in response to human gaze. In this paper, we elaborate on the psychophysics method used in our experiment, which can be a valuable addition to the repertoire of HRI techniques, and we discuss how our findings relate to the field of HRI.

2. COUNTERPREDICTIVE CUEING

Psychophysicists interested in low-level, reflexive visual attention have developed an experimental method for determining the general location of attention in a visual scene [3]. Called *counterpredictive cueing*, this method is based on the idea that people will respond more quickly to a visual cue that is in their field of attention than to a cue that is not.

In a traditional cueing paradigm, participants are asked to watch an on-screen display and to press a button when a visual probe appears. Response times from appearance of probes to button presses should be shorter for probes that appear in an area of the display that is being attended to. Therefore, comparing the response times of visual probes appearing in different areas of the display when the same stimulus is presented can reveal where attention is directed in response to that stimulus. People shift attention in response to directional stimuli at as early as three months of age [5].

When a centrally located stimulus points away from the probe's location (counterpredictive cueing), however, social stimuli such as faces and eyes have different effects than non-social stimuli like arrows. Faces and eyes cue attention in the direction they are pointed, even when participants are motived to attend in the opposite direction, but arrows and extended tongues do not [2, 3, 4]. Functional MRI studies show activation of different cortical pathways for the same stimulus presented either as eyes or a non-social directional symbol [6]. In short, social symbols like eyes and faces seem to cue reflexive attention shifts, while non-social directional symbols like arrows exhibit weak or no reflexive attentional influence; these differences may stem from the existence of face-specific cognitive pathways.

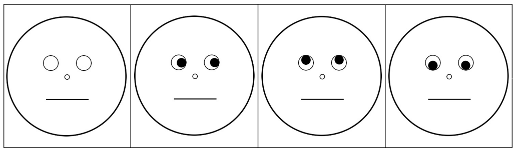
In this experiment, we followed an established counterpredictive cueing paradigm [4] to test the effect of robot faces on reflexive attention. We were interested in whether robot faces would be seen as social entities, and therefore cue reflexive attention, or as non-social directional symbols, which would not be susceptible to counterpredictive gaze.

3. ROBOT GAZE DIFFERS FROM HUMAN GAZE

Details of this experiment have been previously published

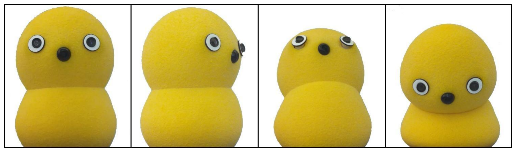
Proceedings of the "Gaze in Human-Robot Interaction" Workshop held at the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012). March 5, 2012, Boston, MA, USA.

(a) Human gaze stimulus

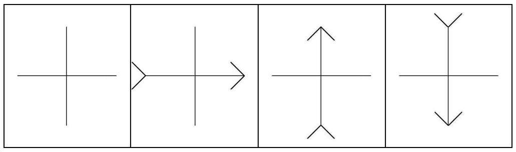


(b) Line-drawn face stimulus

(c) Zeno gaze stimulus



(d) Keepon gaze stimulus



(e) Arrow stimulus

Figure 1: Each subject was assigned to one of five stimulus conditions. This figure shows the front, right, up and down versions of each stimulus.

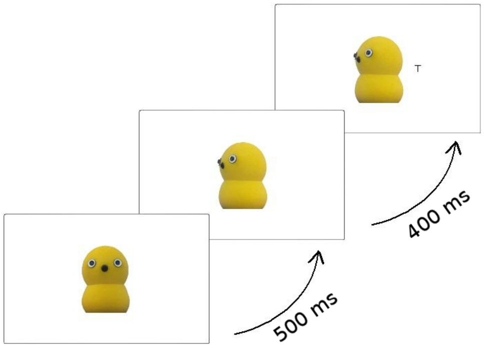


Figure 2: Time course for a single (predicted) trial of the Keepon gaze condition. Setup is similar for other stimuli and directions.

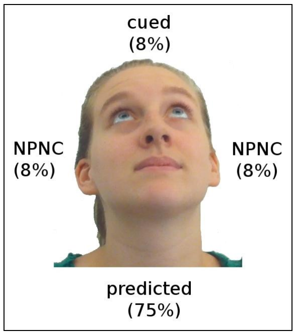


Figure 3: Three types of trials were presented: *cued* (probe and gaze are congruent), *predicted* (probe is in opposite direction to gaze), and *not-predicted-not-cued* or *NPNC* (probe is on a different axis to gaze). Percentages indicate probability of occurrence.

[1], but will be summarized here.

Stimuli for this experiment are shown in Figure 1: a photograph of a human face, a line drawing of a human face, a photograph of a humanoid robot named Zeno, a photograph of a more stylized robot named Keepon, and a line drawing of an arrow. Participants saw a single stimulus type throughout the experiment.

Each trial presented the front-facing stimulus for 500ms, followed by a side-facing stimulus (either left, right, up or down) for 400ms, followed by the appearance of a probe letter (T or L) either to the left, right, above or below the stimulus image (Figure 2). Participants were asked to press the keyboard key corresponding to the probe letter as quickly as possible after the probe appeared.

Seventy participants each viewed 96 trials of this probe response task. Three quarters of the trials were counterpredictive: the stimulus appeared opposite the direction of gaze or arrow, in what we call the “predicted” location. Probes appeared in the other three locations with equal probability, approximately 8% of the time (Figure 3). The location in the direction of gaze or point is called “cued,” and the two locations that are neither in the direction of the stimulus nor exactly opposite the stimulus are called “not-predicted-not-cued” or “NPNC.” Participants were told about this distribution, so they were motivated to shift their attention opposite the direction indicated by the stimulus.

Analysis of response times (RTs) for correct probe presses revealed that non-robot stimuli elicited the reflexive cueing effect identified previously, but that neither robot’s gaze caused the cueing effect. As shown in Figure 4, RTs were statistically shorter for predicted (i.e., opposite direction of gaze or arrow) than for baseline not-predicted-not-cued trials in all stimulus conditions, indicating that participants were attending more frequently to predicted locations than to NPNC locations, as expected. For both Zeno and Keepon stimuli, however, RTs were also statistically shorter on predicted trials than on cued trials, indicating that participants were attending significantly more frequently to the predicted location—opposite gaze—than to the cued location in the direction of gaze. As expected, this difference between predicted and cued trials was absent with the human face stimuli, and was even absent with the arrow stimulus. In other words, when viewing robot faces, people were able to overcome any reflex to attend in the direction of the robot’s gaze, and instead looked opposite that gaze to the predicted location with significantly greater frequency. On the other hand, for human faces and arrows, participants reflexively oriented their attention toward and away from the stimulus direction at statistically indistinguishable rates.

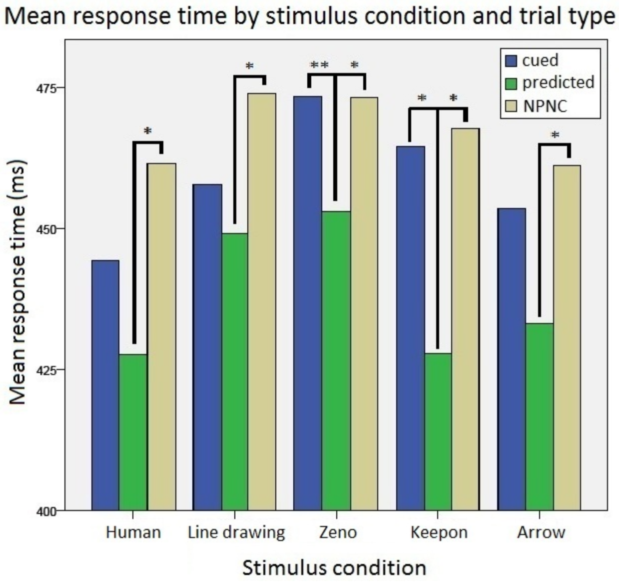


Figure 4: Mean response times in milliseconds for each trial type (cued, predicted and NPNC) and stimulus condition. A single asterisk indicates significant differences ($p < 0.05$), a double asterisk indicates borderline significant differences ($p < 0.10$).

4. DISCUSSION

Previous HRI research has found that robot gaze is a social cue. Our research adds another dimension to this knowledge with the finding that robots are processed differently from other social cues on a perceptual level. Psychology has established that there may be specialized cognitive pathways for processing faces and eyes, but our research suggests that robots are not sufficiently "human-like" to trigger these pathways on a very short timescale, although robots do cue higher-level responses to gaze that may exist outside this face-specific pathway. We can say that robot gaze is meaningful and social, but it is important that we recognize the distinction between low-level perceptual processing and high-level cognitive processing when we do so. Visual perception of robots is multifaceted; understanding how robots do and do not cue social responses is important for designing social robots and for planning human-robot interactions.

There are several possibilities to explain the discrepancy between high-level and low-level perceptions of robots. First, participants in each condition were told which stimulus they were going to see. It is possible that knowing they were looking at a robot encouraged participants to view the gaze as less meaningful; knowing they were looking at a picture or line drawing of a human face made the gaze more meaningful. Anecdotally, participants viewed Zeno as more human-like than Keepon—some participants asked after the experiment whether the Zeno images were of a robot or of a human face—but both Zeno and Keepon failed to cue reflexive attention shifts with their gaze. This difference in category—human versus robotic—could also explain why both robots had similar non-cueing effects, despite differences in appearance. Future experiments could control for context by not identifying the stimulus participants will see, or by presenting multiple stimuli in random order to each participant.

Another possibility is that participants' non-familiarity with robots might have affected how quickly they processed the robot's visual appearance and therefore how effective the reflexive cue could be. In general, people can process gaze information, such as direction, very quickly. Results of this experiment show that participants took longer on average to respond to probes when viewing either robot stimulus, suggesting that the additional processing time necessary to extract gaze information from the unfamiliar robot faces overcame any reflexive cue that might have existed. Interestingly, the more anthropomorphic Zeno seemed to elicit slightly longer response times than the visually stylized Keepon. This may be an effect of Keepon's visual simplicity.

Using psychophysics or other methods, future experiments can further explore how robot gaze affects human attention. Eye-tracking studies, for instance, could reveal more depth about where people are attending during this kind of fast processing task. Cueing attention through gaze is an important element of robot social interactions, and future work is needed to explore this rich but young area.

5. ACKNOWLEDGMENTS

Thank you to two anonymous reviewers for their helpful comments. This material is based upon work supported by grants from Microsoft Research and the National Science Foundation under contracts No. 0835767, No. 0968538, and No. 1117801, and an NSF Graduate Research Fellowship.

6. REFERENCES

[1] H. Admoni, C. Bank, J. Tan, M. Toneva, and B. Scassellati. Robot gaze does not reflexively cue human attention. In *Proceedings of the 33rd Annual conference of the Cognitive Science Society*, pages 1983–1988, 2011.

[2] P. E. Downing, C. M. Dodds, and D. Bray. Why does the gaze of others direct visual attention. *Visual Cognition*, 11(1):71–79, 2004.

[3] J. Driver, G. Davis, P. Ricciardelli, P. Kidd, E. Maxwell, and S. Baron-Cohen. Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6(5):509–540, 1999.

[4] C. K. Friesen, J. Ristic, and A. Kingstone. Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2):319–329, 2004.

[5] B. M. Hood, J. D. Willen, and J. Driver. Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, 9(2), March 1998.

[6] A. Kingstone, C. Tipper, J. Ristic, and E. Ngan. The eyes have it!: An fMRI investigation. *Brain and Cognition*, 55:269–271, 2004.

[7] B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita. Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Human Robot Interactions (HRI'09)*, pages 61–68, La Jolla, California, USA, March 2009. ACM.

[8] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita. Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In *Human Robot Interactions (HRI'09)*, La Jolla, California, March 2009. ACM.