Are You Looking At Me?
Perception of Robot Attention is Mediated by Gaze Type and Group Size

Henny Admoni, Bradley Hayes, David Feil-Seifer,
Daniel Ullman, and Brian Scassellati
Department of Computer Science, Yale University
New Haven, Connecticut 06520 USA
{henny.admoni, bradley.h.hayes, david.feil-seifer, daniel.ullman, brian.scassellati}@yale.edu

Abstract—Studies in HRI have shown that people follow and understand robot gaze. However, only a few studies to date have examined the time-course of a meaningful robot gaze, and none have directly investigated what type of gaze is best for eliciting the perception of attention. This paper investigates two types of gaze behaviors—short, frequent glances and long, less frequent stares—to find which behavior is better at conveying a robot's visual attention. We describe the development of a programmable research platform from MyKeepon toys, and the use of these programmable robots to examine the effects of gaze type and group size on the perception of attention. In our experiment, participants viewed a group of MyKeepon robots executing random motions, occasionally fixating on various points in the room or directly on the participant. We varied type of gaze fixations within participants and group size between participants. Results show that people are more accurate at recognizing shorter, more frequent fixations than longer, less frequent ones, and that their performance improves as group size decreases. From these results, we conclude that multiple short gazes are preferable for indicating attention over one long gaze, and that the visual search for robot attention is susceptible to group size effects.

Index Terms—gaze; group dynamics; social robotics; human-robot interaction

I. Introduction

Eye gaze is a critical component of typical social interactions. We use gaze to indicate attention, whether toward a speaker or toward an object of mutual interest. However, subtle gaze timing can have a strong effect on realism and comfort in an interaction. Gaze fixations that are too short can be interpreted as shyness, avoidance, or disinterest. Gaze fixations that last too long can appear menacing or uncomfortable. With the development of real-world robotic systems comes a need to understand and use gaze cues effectively.

Human-human conversation partners frequently direct their gaze toward the person to whom they are listening or speaking [1], [2], using mutual gaze to signify attention. Robot gaze seems to be leveraged just as well as human gaze; for example, people use both human gaze [3] and robot gaze [4] to successfully disambiguate referential utterances.

Many HRI gaze studies use what we call a *behavioral* level of analysis: measurements take the form of observable behaviors and explicit self-reports. In HRI, interactions tend to occur between one robot and one human (e.g., [5]–[7]) or between one robot and a small human group (e.g., [8]). Because modern-day social robots tend to be individual machines operating amidst a group of humans, this level of analysis addresses observable human-robot interactions in typical HRI environments. Such research has identified the general effectiveness of coherent robot gaze, for instance, in cueing conversational roles [8] and improving recall of stories [6].

Other gaze studies in HRI investigate robot gaze at a lower level of analysis, which we call the *perceptual* level. These studies tend to be psychophysical in nature, measuring millisecond-level reflex responses rather than broad behaviors or reported opinions. One such study found that robot gaze does not seem to cue reflexive shifts of attention the way human gaze does [9], suggesting that robot and human stimuli are processed differently in early visual pathways.

In the current work, we seek to understand which features of a robot's gaze make that robot appear to be attending to someone. There are many components of gaze behavior: frequency, duration, and locations of fixations; scan paths taken to reach fixation points; congruency of fixations during mutual gaze and joint attention. Because making eye contact is a strong signifier of attention, in this study we fix the gaze location on the participant, and manipulate duration and frequency of fixations. We contrast gaze behaviors along a spectrum, from short, frequent glances to longer, less frequent stares, all directed at the user.

To quantitatively examine the effects of these two gaze types on the perception of attention, we measure the detection rate of attention fixations (i.e., fixations directed at the user) over three conditions on the spectrum of gaze types. To measure the detection rate of fixations, we present the target (the robot displaying attention fixations) among a number of distractors (identical robots displaying fixations not directed at the user). This target-amidst-distractors method is common in psychophysical studies for determining a detection level, though it is (currently) uncommon for HRI experiments. We call this a *sociophysical* level of analysis because the measurement techniques (accuracy of identification) and methods

(target-amidst-distractors) is drawn from psychophysics, but the task (identifying attention through gaze) is a social one. In addition to gaze type, we also manipulate group size to identify whether the number of distractors has an effect on participants' ability to recognize attention. This experiment is described in Section IV.

Our search for an effective robot platform that could be used in a multi-robot experiment led us to create programmable research tools out of a readily available toy called MyKeepon. Section III describes the hardware and software modifications we used to create this novel research platform.

II. Related Work

Gaze recognition develops early and remains critical for non-verbal communication throughout life. Newborns [10] and older infants [11] show preferences for open eyes over closed ones. Adults are highly accurate at detecting another person's face-directed gaze during normal conversations [1]. In four-person conversations, researchers found an 88% probability that the person being looked at is the one being listened to, and a 77% probability that the person being looked at is the one being spoken to [2]. Eye gaze is also a useful cue in disambiguating referential expressions in dialogue. In an experiment where conversation partners verbally referenced objects on their displays, participants successfully used gaze cues to distinguish between competing referents before the linguistic point of disambiguation [3]. Interestingly, people tend to overestimate the amount of gaze directed at their own faces, mistaking a look over their shoulder for a gaze to their face [1].

In visual search tasks involving the selection of one unique item from among a group of distractors, participants were faster at detecting eyes gazing straight ahead from among left-gaze and right-gaze distractors than they were at detecting left-gazing eyes from right-gazing and straight ahead distractors, or detecting right-gazing eyes from among left-gazing and straight ahead distractors [12]. This effect is maintained even when the stimuli are detail-impoverished schematic representations of eyes, but disappears when the stimuli are geometric shapes instead of eyes. Mutual eye gaze also leads to faster processing, such as categorization of gender and access to semantic knowledge, than averted gaze [13].

Eye tracking studies reveal that gaze is affected by context. Head-mounted eye trackers show that gaze is task-driven, and that fixation locations are determined by the task at hand and learned over time [14]. Dual eye tracking has shown that the occurrence of mutual gaze, where two partners look at each other, depends on the dynamic interplay of behaviors and characteristics of both partners [15].

Functional MRI studies identify differences for processing different features of gaze. Gaze duration is processed in the medial prefrontal cortex, an area that is involved with more complex metacognitive processing, which is distinct from the brain region processing gaze direction [16], [17]. In other words, gaze duration is a distinct feature which is processed independently of other gaze features. The intensity of brain activity in response to gaze shifts is modulated by context; fMRI studies show that activity in the superior temporal sulcus is affected by whether a virtual agent correctly or incorrectly shifts its gaze toward a target [16], [18].

Several robotics researchers have explored how robot gaze influences human-robot interactions. A number of studies have tried to improve human-agent communication through appropriate gaze, both in robotic systems [5], [6], [8], [19]–[22] and in virtual intelligent agents [23]–[25]. Gaze cues can influence human participants in a human-robot interaction to conform to intended conversational roles such as addressee, bystander, and nonparticipant [8]. People are also better at recalling details of a story when the robot storyteller gazed at them more frequently [6]. A robot that responds to and maintains joint attention improves task performance and receives higher ratings for competence and social interactivity than a robot that does not display joint attention behaviors [5]. Unlike eye gaze, however, people are sensitive to a robot's direct gaze but not to a nearby indirect gaze [26].

Using eye tracking, researchers found that participants follow a robot's gaze, even when the task does not require them to do so [4]. They also found that when a robot's gaze and utterances are congruent, participants can judge utterances more quickly than when gaze and utterances are incongruent. On the other hand, when examining millisecond-level psychophysical responses, robot gaze does not cue the same reflexive attention shifts that human gaze does, instead seeming to be susceptible to top-down control [9].

In this paper, we are interested in how gaze frequency and duration affects the perception of attention. Some previous work attempts to specifically investigate these features of gaze during interactions. One such study found that a speaker looked at the face of an addressee between 25% and 56% of the time, depending on how many other people were involved in the conversation [8]. Researchers found that gaze switch timings consistent with human timings appeared more natural than gaze switches that occurred with every speech utterance [20]. Too much gaze was also a problem, however: high levels of mutual attention without valence or responsiveness decreased rapport with a virtual agent [27].

Research in joint attention has also investigated gaze timings. One study found that a person's gaze dwelled on a referenced object for approximately 1.9 seconds on average, with no statistical difference in the amount of time spent looking when the referencer was human or a virtual agent [25]. Another such data-driven study of micro-level behaviors found that participants look at a communication partner's face (whether human or agent) within about 800 to 900 milliseconds after their partner's head movements and 600 to 700 milliseconds after naming an object for their partner to learn [7]. Participants spent longer fixating on a robot partner's face than a human partner's face, however.

III. Programming MyKeepon

In order to examine the effects of gaze duration on the perception of attention, we sought to use a robot platform with

highly salient visual features (e.g., eyes) with an otherwise simple appearance. Keepon is a small, yellow, snowman-like robot with two eyes and a nose, but no other facial features. Originally designed for applications such as autism therapy, Keepon is a socially evocative robot that has been shown to elicit various social behaviors from children and adults [28]. The original research-grade robot is easy to control but expensive to buy, making it infeasible to use in our current study, which requires multiple robots. Fortunately, a version of Keepon is available as an inexpensive consumer-grade toy under the name MyKeepon from BeatBots LLC. In this section, we describe how we converted MyKeepon toys into programmable research tools. For more details and photographs of the process, please see our website at http://hennyadmoni.com/keepon/.

MyKeepon has four degrees of freedom (DOFs) using three DC motors. It can lean forward and backward, lean left and right, rotate clockwise and counterclockwise on its base, and bob up and down. For this project, we number the motors arbitrarily: motor one controls rotation on the base, motor two controls left/right lean and bob, and motor three controls forward/back lean. Motor two's control is switched between its two DOFs using a small geared rocker mechanism; we found this mechanism difficult to control and therefore we only employ motors one and three in this experiment.

In the toy version of MyKeepon, motors are controlled by an internal circuit board. To take control of the robot's motors, we circumvented the internal board and soldered wires directly to the leads of each motor. We removed MyKeepon's internal control board along with the microphone, speakers and battery housing.

We use Arduino, an open-source hardware platform, as a control interface from computer to robot motors [29]. Each MyKeepon robot is attached to one Arduino Uno and one Arduino Motor Shield, which plugs into the Arduino Uno and is designed to run up to four DC motors. Each Arduino Uno is connected to the computer through its USB connector; when controlling multiple robots, we use a USB hub between the Uno boards and the computer.

The USB connection to the Arduino Uno allows us to send commands from the computer to the motors over a serial connection. To ensure replicability between participants, robot motions are pre-scripted, though they can be calculated and sent in real time. Each robot's motions are designed on the computer, then sent at the appropriate time to the Arduino Uno board attached to the robot. Commands are cached on the board until execution time, at which point the commands are played back sequentially, causing the motors (and the robot) to move. Figure 1 shows the hardware setup with control computer, USB hub, Arduino Uno and Motor Shield pairs, and MyKeepon robots. Though only three robots are shown in this figure, the setup is similar for any number of robots.

The simple DC motors in MyKeepon robots are less sophisticated than typical high-precision motors used in research, specifically in the absence of encoders to report precise positioning. We compensated for this limitation with hand-tuning
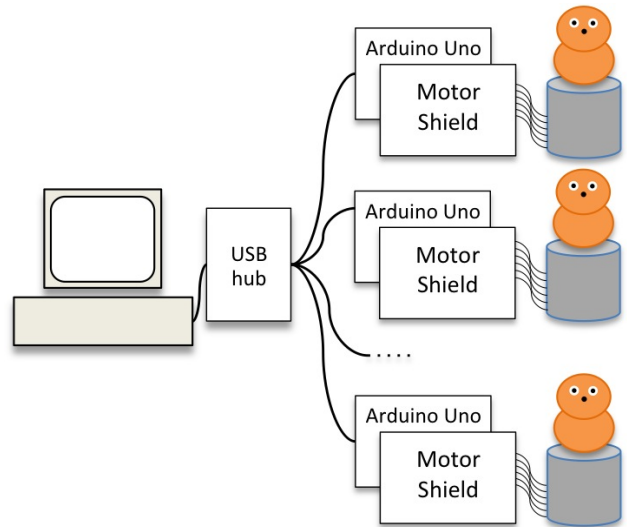

Fig. 1. A diagram of the hardware setup. Robot motors were wired to an Arduino Motor Shield, which paired with an Arduino Uno to receive commands from a computer via USB.

when necessary, but these motors are the major limitation for using MyKeepon as a research platform.

Several pieces of code design, transmit, and control robot motions, some running on the computer and others running on the Arduino boards. Computer-based code includes a movement generator that automatically designs robot motions given some criteria, such as direction and duration of gaze fixation. A Python script interfaces with code running on the Arduino by sending movement commands over a serial connection via USB. For instance, the Python command

1: move(keepon_ID, motor_num, time)

moves *motor_num* for *time* milliseconds on robot with ID number *keepon_ID*. The robot's ID number is hardcoded on its Arduino board.

We use a publicly-available package called AFMotor to control the DC motors in MyKeepon from the Arduino. To communicate with this low-level control, we wrote a state-based controller that listens for move commands arriving at the serial port and issues appropriate calls to AFMotor.

IV. EXPERIMENT

We conducted an experiment using these programmable MyKeepons to evaluate the effects of gaze duration and group size on the perception of attention. The experiment was a mixed 3 (group size) x 4 (gaze duration) between- and within-subjects design. Participants viewed a group of robots (four, six, or eight, between-subjects) making simultaneous random motions. Between random movements, each robot occasionally fixated its gaze on various positions in the room for a given duration (zero, one, three, or six seconds, within-subjects); during these occasional fixations, a specific robot (the *target* for that trial) always fixated on or near the participant. All robots fixated for the same duration in a single trial and the total duration of fixation was held constant among trials; robots

fixated six times on a one-second fixation trial, twice on a three-second fixation trial, and once on a six-second fixation trial. This inverse relationship between duration and frequency evokes the appearance of different gaze types, from frequent brief glances to longer stares. Each robot was the target in an approximately equal number of trials. After each trial, participants recorded which robot they thought was paying attention to them, as well as their confidence in that decision.

Our hypotheses are as follows:

H1   The type of gaze fixation affects accuracy: multiple short glances will be easier to detect than fewer longer fixations.

H2   The size of the group affects accuracy: more distractor robots will make it harder to detect the gaze of the target robot.

MyKeepon motor control is somewhat imprecise, so perfectly direct gaze toward participants is difficult to achieve. Each robot's movements were hand-calibrated to assure fixation toward the participant's location, though assuring that target robots directly oriented toward participants was challenging. In the experiment we report below, target robots fixated on or near participants on the target trials. Though robot fixations were not as precise as human fixations, this only served to make the task more challenging and to strengthen the results. Despite their imprecision, robot motors tend to be consistent, so whatever errors were present in target fixations likely existed for all participants.

A. Apparatus

MyKeepon robots were placed side-by-side in a containment apparatus which was covered in a black cloth (Figure 2). The apparatus was approximately 152cm wide by 61cm deep by 15cm tall. The robots were placed side-by-side with about 20cm from the center of one robot base to another. Figure 3 shows an overhead schematic of the experiment setup.

In the six robot condition, the two outermost robots were removed, and in the four robot condition the four outermost robots were removed, so that the robots present were always centered within the apparatus. Colored labels on the front of the box are used during the experiment to refer to robots. We chose not to use numbers in an attempt to avoid ordinal effects.

Participants were seated about 152cm away from and centered on the midpoint of the apparatus. At this distance, the total robot display subtended approximately 25° of the participant's visual field in the 4-robot condition (69 cm), 38° in the 6-robot condition (104 cm), and 49° in the 8-robot condition (140 cm). Each robot (9 cm across) subtended 3.4° of the visual field, and an individual robot eye (1.3 cm) subtended approximately 1° of the visual field. Although the size of an individual robot's eye is quite small, the eyes are fixed to its body, so the robot moves its entire body to orient its gaze. In every condition, participants needed to move their eyes and possibly their heads to foveate on every robot.
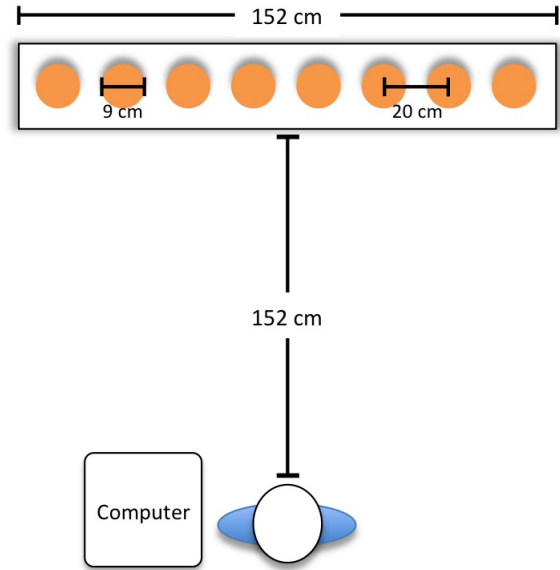

Fig. 3.   Overhead schematic of the experiment setup showing robots, containment apparatus, computer and participant. This figure is not to scale.

B. Procedure

Fifty-three participants (20 male, 33 female) took part in the experiment. The experiment took approximately 30 minutes, and participants were paid for their time. Each participant was randomly assigned to one group size condition (four, six, or eight robots).

Each participant viewed 30 trials, and each trial was comprised of 30 seconds of pre-scripted movement. In each trial, the robots exhibited automatically generated random motions in two DOFs, leaning forward or back and rotating clockwise or counterclockwise on their bases. The two DOFs could move simultaneously, causing the robots to appear to be looking around the room. At approximately equally spaced intervals, but not necessarily simultaneously, each robot stopped its motion for a set amount of time before returning to performing random motions; we call this a *gaze fixation*, and the apparent location toward which the robot is oriented is its *fixation location*. For the target robot in a trial, the fixation location was always the participant; other robots oriented toward various points in the room during their fixations. For example, in a three-second fixation trial, the target robot fixated on the user and distractor robots fixated on various locations approximately every ten seconds, though the fixation periods did not necessarily overlap. Because behaviors between fixations were random, the robots had to move different distances from different positions to return to their fixation locations throughout the trial. Trial presentation order was randomized across participants.

Robots fixated for zero, one, three, or six seconds per trial. Zero-second fixations were a control, in which robots did not stop their random movements. One-second fixations were selected based on preliminary testing, which revealed that one-second fixations are brief enough to be difficult to identify in
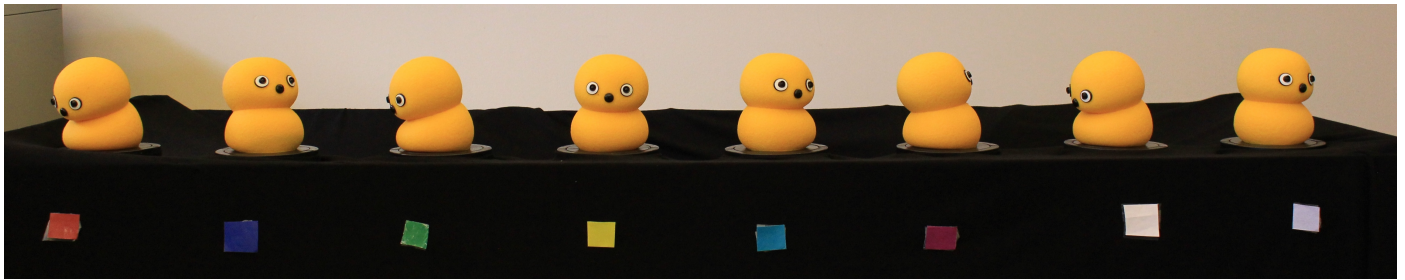
Fig. 2. A photograph of a participant's view of the eight robot condition. The fourth robot from the left (with a yellow label) is fixating on the participant; the other robots are gazing elsewhere.

the eight robot condition. Six second fixations were chosen to be easily recognized, and three seconds was chosen as easily divisible to maintain total fixation duration in a trial. There were six zero-fixation trials and eight of each other fixation duration for a total of 30 trials.

Participants were seated next to a computer, which recorded their results and controlled the robots. Participants began a trial by clicking a "Start" button displayed on the computer monitor, which initiated the robots' pre-scripted movements for that trial. At the conclusion of each 30 second trial, a screen appeared on which participants selected which robot they believed was paying attention to them. They assigned a confidence value (from 0 to 100) to their choice by using a slider bar with whole-number increments. If they were able to make a decision sooner, participants could press the "Enter" key on the computer's keyboard to bring up the selection screen immediately, though the robots continued to move for the full 30-second trial. Before data recording began, participants engaged in two practice trials under experimenter supervision.

V. RESULTS

Two participants were excluded from analysis due to technical malfunction. Additionally, four individual trials were excluded due to failure to respond or error in recording a response. We analyzed the results of 51 participants (25 in the eight-robot group, 19 in the six-robot group, and 7 in the four-robot group), for a total of 406 trials of one-second fixations, 408 trials of three-second fixations, 406 trials of six-second fixations and 306 trials of zero-second fixations across all robot group sizes. Figure 4 shows average accuracy for each fixation duration as a function of group size.

We conducted a mixed-model repeated measures ANOVA with fixation duration (0, 1, 3, or 6 seconds) as the within-subjects repeated variable and group size (4, 6, or 8 robots) as the between-subjects variable. There is significance for fixation time ($F(3, 144) = 17.503, p < 0.001$) and group size ($F(2, 48) = 5.105, p = 0.010$). There is also a significant interaction effect ($F(6, 144) = 3.554, p = 0.003$).

Pairwise comparisons of fixation duration reveal that one- and three-second fixations led to significantly higher accuracy than either zero- or six-second fixations ($p \leq 0.001$ in all cases). There was no statistical difference in accuracy
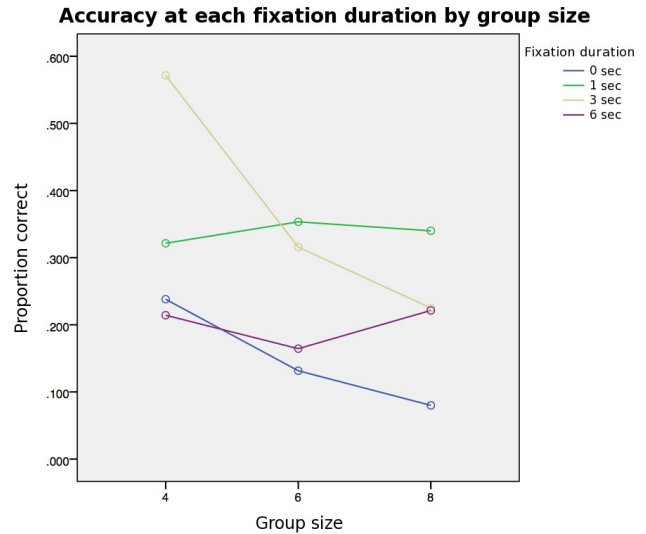

Fig. 4. Proportion of accurate responses as a function of group size and fixation duration. The line for zero-second fixations is at approximately chance.

between zero- and six-second fixations, or between one- and three-second fixations. Given that zero-second fixations had accuracy of about chance for each group size (0.24 for four robots (chance is 0.25), 0.13 for six robots (chance is 0.17), 0.08 for eight robots (chance is 0.13)), six-second fixation accuracies were not statistically different than chance, though one- and three-second fixation accuracies were significantly better than chance.

Post-hoc analysis of group size using a Bonferroni correction found that accuracy in the four robot group was significantly better than accuracy in the eight robot group ($p = 0.007$) and marginally better than accuracy in the six robot group ($p = 0.054$). No statistical difference was found for accuracy between six and eight robot groups.

VI. DISCUSSION

MyKeepon robot motors are imprecise but consistent, so the target robot's gaze was offset by some small amount from a perfectly direct gaze on many trials. People overestimate the amount of gaze directed toward their face [1], but even so, this gaze offset may explain generally low accuracy rates. On the other hand, the motors appear to be consistent across many

trials, so we are confident that stimuli were consistent across presentations. We present our results with the understanding that fixation errors are higher with robots as stimuli than they would be with humans.

H1 predicted that accuracy would improve as gaze fixation changed from long and infrequent to short and frequent. Results support this, with statistically significant differences in accuracy between short (one- and three-second) and long (six-second) fixations. Since total fixation time in a single trial was held constant, this suggests that multiple short fixations are better at conveying attention than fewer longer fixations. We predict that it is the transition from motion to gaze fixation, rather than the fixation itself, that cues the perception of attention. Therefore, people may be responding to "fixation events"—the transition between movement and fixation— rather than to active gaze. There were six times as many fixation events in the one-second condition than in the six-second condition, perhaps accounting for improved accuracy on shorter fixation trials. If our prediction is correct, this would be an interesting finding about human gaze processing.

H2 predicted that group size would have a negative effect on accuracy, which was also supported. Results show significant differences in accuracy rates between the four-robot group and the other groups, with accuracy rates of 37%, 28%, and 26% over all fixation durations, respectively. This is consistent with findings from a visual search task, where more distractors led to a degradation in performance when detecting straight eye gaze [12]. These results suggest that eye gaze, while an important social factor, does not cause a "pop out" effect like some more basic stimuli.

Meaningful eye gaze consists of many features in addition to frequency and duration of fixations. For example, the velocity of a saccade and the scan path also reveal socially relevant information. Furthermore, different behaviors such as gaze following, joint attention, or attention maintenance may require the use of different features of gaze. The current experiment breaks down the complexity of social gaze by isolating frequency and duration in the context of indicating attention. A full exploration of gaze necessitates understanding all the features of gaze and their interplay, and is a rich avenue for future research.

One difference between this and most other HRI studies is the use of a multi-robot setup. Human-human and human-robot interactions outside of the laboratory do not occur in a vacuum. There are competing visual stimuli in many real-world tasks that draw attention away from a visual target. In multi-robot domains, gaze features combine not only within a single robot's behaviors but across robots, making for a complex visual scene. To make progress toward a more holistic understanding of HRI, we must continue to explore visual attention under distracting and difficult conditions.

As with many lab-based experiments, we must consider how transferable the findings of our study are to real-world interactions. The current work is a necessary step, but not a final point, on the path to understanding natural gaze. Given that, our work yields some suggestions for the design of robots and robot behaviors in HRI. Because frequent short glances were more easily recognized in this experiment, looking at a user to initiate or maintain an interaction may be most effective using short, frequent glances, rather than an extended stare. This is supported by research suggesting that an agent that maintains mutual gaze for an extended duration (without other social gestures) leads to strongly negative responses from users [27], but that gazes that are too short and frequent also hinder communication [20]. It would be interesting to identify whether this short-and-frequent gaze preference is also present in other gaze scenarios like joint attention, where gaze is directed toward an object of mutual interest, rather than at the user themselves. Because context plays a role in the control of eye gaze [16], [18], an experiment that tests gaze features during task performance (like providing driving directions) might reveal different features at work in signifying attention.

VII. Conclusion

Robot gaze has been shown to be successful at indicating attention to people or objects, but gaze duration and frequency affect that perception of attention. We investigated the effects of gaze type (short, frequent glances versus long, infrequent stares) and group size on the detection of attention using programmable MyKeepon robots, which were devised and built for this project. We ran an experiment in which groups of four, six, or eight MyKeepon robots performing random movements occasionally fixated their gaze on either the participant or on other locations. We evaluated participants' accuracy in detecting these fixations for gaze durations of zero, one, three, and six seconds. We found that multiple, short fixations are better at conveying attention than fewer, longer gazes. More distractors made this task more difficult, with accuracy decreasing for larger group sizes. Our results have implications for robot designers who want to make their robots appear to be attending to users, as well as for psychologists who want to understand gaze.

References

[1] M. Argyle and M. Cook, *Gaze and Mutual Gaze*. Oxford, England: Cambridge University Press, 1976.

[2] R. Vertegaal, R. Slagter, G. van der Veer, and A. Nijholt, "Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY USA: ACM, 2001, pp. 301–308.

[3] J. E. Hanna and S. E. Brennan, "Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation," *Journal of Memory and Language*, vol. 57, pp. 596–615, 2007.

[4] M. Staudte and M. W. Crocker, "Investigating joint attention mechanisms through spoken human-robot interaction," *Cognition*, vol. 120, pp. 268–291, 2011.

[5] C.-M. Huang and A. L. Thomaz, "Effects of responding to, initiating and ensuring joint attention in human-robot interaction," in *20th IEEE International Symposium on Robot and Human Interactive Communication (2011 RO-MAN)*, Atlanta, GA USA, 2011, pp. 65–71.

[6] B. Mutlu, J. Forlizzi, and J. Hodgins, "A storytelling robot: Modeling and evaluation of human-like gaze behavior," in *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids '06)*, 2006, pp. 518–523.

[7] C. Yu, P. Schermerhorn, and M. Scheutz, "Adaptive eye gaze patterns in interactions with human and artificial agents," *ACM Transactions on Interactive Intelligent Systems*, vol. 1, no. 2, January 2012.

[8] B. Mutlu, T. Kanda, J. Forlizzi, J. Hodgins, and H. Ishiguro, "Conversational gaze mechansims for humanlike robots," *ACM Transactions on Interactive Intelligent Systems*, vol. 1, no. 2, January 2012.

[9] H. Admoni, C. Bank, J. Tan, and M. Toneva, "Robot gaze does not reflexively cue human attention," in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society (CogSci 2011)*, L. Carlson, C. Hölscher, and T. Shipley, Eds. Austin, TX USA: Cognitive Science Society, 2011, pp. 1983–1988.

[10] A. Batki, S. Baron-Cohen, S. Wheelwright, J. Connellan, and J. Ahluwalia, "Is there an innate gaze module? Evidence from human neonates," *Infant Behavior and Development*, vol. 23, pp. 223–229, 2000.

[11] R. Brooks and A. N. Meltzoff, "The importance of eyes: How infants interpret adult looking behavior," *Developmental Psychology*, vol. 38, no. 6, pp. 958–966, 2002.

[12] M. von Grünau and C. Anston, "The detection of gaze direction: A stare-in-the-crowd effect," *Perception*, vol. 24, pp. 1297–1313, 1995.

[13] C. N. Macrae, B. M. Hood, A. B. Milne, A. C. Rowe, and M. F. Mason, "Are you looking at me? eye gaze and person perception," *Psychological Science*, vol. 13, no. 5, pp. 460–464, 2002.

[14] M. Hayhoe and D. Ballard, "Eye movements in natural behavior," *Trends in Cognitive Sciences*, vol. 9, no. 4, pp. 188–194, 2005.

[15] F. Broz, H. Lehmann, C. L. Nehaniv, and K. Dautenhahn, "Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation," in *Proceedings of the "Gaze in Human-Robot Interaction" Workshop at the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012)*, Boston, MA, USA, March 2012.

[16] D. Bristow, G. Rees, and C. D. Frith, "Social interaction modifies neural response to gaze shifts," *Social Cognitive and Affective Neuroscience*, vol. 2, pp. 52–61, 2007.

[17] B. Kuzmanovic, A. L. Georgescu, S. B. Eickhoff, N. J. Shah, G. Bente, G. R. Fink, and K. Vogeley, "Duration matters: Dissociating neural correlates of detection and evaluation of social gaze," *NeuroImage*, vol. 46, no. 4, pp. 1154–1163, 2009.

[18] K. A. Pelphrey, J. D. Singerman, T. Allison, and G. McCarthy, "Brain activation evoked by perception of gaze shifts: the influence of context," *Neuropsychologia*, vol. 41, no. 2, pp. 156 –170, 2003.

[19] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, "Towards a humanoid museum guide robot that interacts with multiple persons," in *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robotics*, 2005, pp. 418–423.

[20] J. G. Trafton, M. D. Bugajska, B. R. Fransen, and R. M. Ratwani, "Integrating vision and audition within a cognitive architecture to track conversations," in *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, 2008, pp. 201–208.

[21] C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh, "Where to look: A study of human-robot engagement," in *Proceedings of the 9th International Conference on Intelligent User Interfaces (IUI '04)*. New York, NY, USA: ACM, 2004, pp. 78–84.

[22] A. Yamazaki, K. Yamazaki, M. Burdelski, Y. Kuno, and M. Fukushima, "Coordination of verbal and non-verbal actions in human-robot interaction at museums and exhibitions," *Journal of Pragmatics*, vol. 42, pp. 2398–2414, 2010.

[23] J. Cassell, "Embodied conversational interface agents," *Communications of the ACM*, vol. 34, no. 4, April 2000.

[24] E. Gu and N. Badler, "Visual attention and eye gaze during multi-party conversations with distractions," in *Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science, J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Olivier, Eds. Springer Berlin Heidelberg, 2006, vol. 4133, pp. 193–204.

[25] N. Pfeiffer-Lessmann, T. Pfeiffer, and I. Wachsmuth, "An operational model of joint attention - timing of gaze patterns in interactions between humans and a virtual human," in *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012)*, N. Miyake, D. Peebles, and R. P. Cooper, Eds. Austin, TX USA: Cognitive Science Society, 2012, pp. 851–856.

[26] M. Imai, T. Kanda, T. Ono, H. Ishiguro, and K. Mase, "Robot mediated round table: Analysis of the effect of robot's gaze," in *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication*, 2002, pp. 411–416.

[27] N. Wang and J. Gratch, "Don't just stare at me!" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, 2010, pp. 1241–1250.

[28] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keepon: A playful robot for research, therapy, and entertainment," *International Journal of Social Robotics*, vol. 1, pp. 3–18, 2009.

[29] Arduino, http://arduino.cc, September 2012.