Robot Form and Motion Influences Social Attention

Alvin X. Li[1*], Maria Florendo[1], Luke E. Miller[1], Hiroshi Ishiguro[2], Ayse P. Saygin[1]

(1) Cognitive Science, UC San Diego
9500 Gilman Drive, La Jolla, CA
92093-0515
+1-858-822-1994
{axl002, mflorendo, lmiller, asaygin}@ucsd.edu

(2) Department of System Innovation, Graduate school of
Engineering science, Osaka University
Osaka, Japan
+81-6-6850-6360
ishiguro@sys.es.osaka-u.ac.jp

ABSTRACT

For social robots to be successful, they need to be accepted by humans. Human-robot interaction (HRI) researchers are aware of the need to develop the right kinds of robots with appropriate, natural ways for them to interact with humans. However, much of human perception and cognition occurs outside of conscious awareness, and how robotic agents engage these processes is currently unknown. Here, we explored automatic, reflexive social attention, which operates outside of conscious control within a fraction of a second to discover whether and how these processes generalize to agents with varying humanlikeness in their form and motion. Using a social variant of a well-established spatial attention paradigm, we tested whether robotic or human appearance and/or motion influenced an agent's ability to capture or direct implicit social attention. In each trial, either images or videos of agents looking to one side of space (a head turn) were presented to human observers. We measured reaction time to a peripheral target as an index of attentional capture and direction. We found that all agents, regardless of humanlike form or motion, were able to direct spatial attention in the cued direction. However, differences in the form of the agent affected attentional capture, i.e., how quickly the observers could disengage attention from the agent and respond to the target. This effect further interacted with whether the spatial cue (head turn) was presented through static images or videos. Overall whereas reflexive social attention operated in the same manner for human and robot social agents for spatial attentional cueing, robotic appearance, as well as whether the agent was static or moving significantly influenced unconscious attentional capture processes. These studies reveal how unconscious social attentional processes operate when the agent is a human vs. a robot, add novel manipulations to the literature such as the role of visual motion, and provide a link between attention studies in HRI, and decades of research on unconscious social attention in experimental psychology and vision science.

Categories and Subject Descriptors

H.1.2 **[Models and Principles]**: User/Machine Systems – *Human factors.* H.5.2 **[Information Interfaces and Presentation]**: User Interfaces – *Evaluation/methodology, User-Centered Design*

General Terms

Design, Human Factors.

Keywords

Social Attention；Spatial Attention；Humanlikeness；Robot Design；Experimental Psychology

1. INTRODUCTION

Social robots are becoming increasingly prevalent in society, employed in roles such as entertainment, education, and healthcare [1, 2]. Many of these roles, particularly in health care and education, require building of trust and empathy between robots and humans, thus creating a comfortable user experience is important for social robots to be successful. However there are a myriad of issues that remain to be solved to create sociable robots that can reproduce the Human-Human interaction experience. Roboticists must consider issues such as the design of the robot's appearance, how the robot behaves, and how much autonomy the robot possesses, among others.

HRI researchers are well aware of the need to develop the right kinds of robots with appropriate, natural ways for them to interact with humans, and significant progress has been made in recent years in identifying factors of robot design that influence acceptability. However, much of human perceptual and neural processing occurs outside of awareness, and there are many aspects of processing that cannot be measured with observational studies or overt ratings, as has typically been done in prior HRI work. Here, we suggest further insight could be gained by also applying theory and methods from the cognitive sciences that tap into automatic or unconscious social processing that occurs at a millisecond time scale. These studies, while admittedly disembodied from the viewpoint of real life HRI applications, should supplement more naturalistic interaction studies. Developing truly "neuroergonomic" social robotic systems requires an interdisciplinary approach that can benefit from studies that reveal fundamental processes in the human brain that guide social situations, including studies on whether and how these processes generalize to the case of human robot interaction.

In addition to bringing in methods and theory from attention research, in the present study, we examined two aspects of robot design, appearance and motion, and their influence on human social attention. The role of appearance and motion have been of interest to HRI researchers in both experimental work [3, 4] and in theoretical frameworks such as the Uncanny Valley Hypothesis [5, 6]. Attempting to quantify the experience of interacting with a robot is difficult, as it is a complex phenomenon with many internal mental processes at play. In experimental research, most
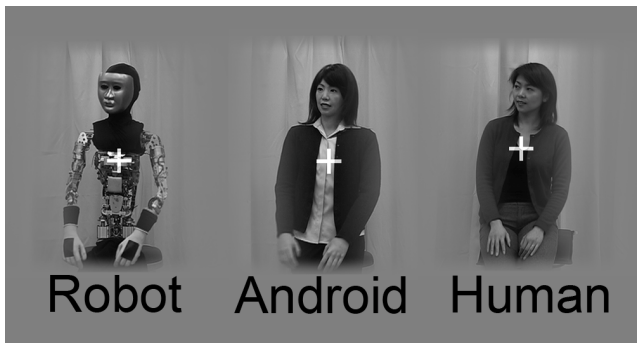
Figure 1: The three agents used in the present study. The agents differ along the dimensions of form and motion. Robot has non-biological form and motion. Android has non-biological form with biological motion. Human has both biological form and motion. Furthermore, subjects were informed that the Robot and Android were machines while the Human was a real person so the agent identity was a difference as well.

prior studies have attempted to characterize the experience of interacting with robots based on subjective judgments and questionnaires. Subjects are asked to rate robots based on categories such as humanlikeness, familiarity, acceptability, sociability, etc. [7-11]. Subjective ratings are important in understanding human reactions to robots, but there are several reasons the field can benefit from complementing these with different experimental approaches and more objective dependent measures. First, there is no consensus as to whether the commonly-used humanlikeness, familiarity, eeriness dimensions are the most suitable for designing more acceptable agents [9, 12]. Second, it is not certain if these rating scales are capable of capturing the subjective experience of interacting with a robot. Finally, subjective ratings have general limitations such as test reliability, test validity, emotional state, and pressure to give socially desirable answers [13].

As a result of these limitations, we should look towards developing more objective approaches and measures to complement questionnaires and surveys in study of HRI. For example, physical approach distance to a robot [14], eye contact [15], eye gaze following [16, 17], dwell time [18], and perceptual adaptation [19] have been used to provide more objective measures for the uncanny valley hypothesis. Neuroimaging methods such as EEG and fMRI can be used to measure how the brain reacts to stimuli of real and artificial agents [20-22]. Objective measures can be modeled with subjective ratings to create a mapping from features of an artificial agent to a behavioral response or neural activity [23]. This eases the process of interpreting how and why a person would have a particular subjective experience when interacting with a robot.

We proposed social attention and the spatial attention task to be a potentially useful objective measure for investigation in HRI. Humans have evolved a sophisticated *Social Attention System* to aid interpersonal interaction and cooperation. This system includes such skills as interpreting facial expressions and understanding non-verbal gestures and actions of others to infer their intent and affective state [24]. In most circumstances this system performs well, allowing us to rapidly and effectively acquire and transmit more communicative and social cues that we recognize. When designing robots that will be immersed in human society, we could benefit from building these machines whilst being cognizant of the social and communicative abilities humans already possess. In other words these machines should move in ways that we can understand [25] and one way to achieve this is to have them use the same social cues we use in everyday communication. Other people's behavior is immensely useful in aiding the detection and location of important social events and objects. Past research in HRI has focused on how effective robots are at drawing and directing overt and conscious social attention [16, 26, 27]. These past studies generally focused on a scenario where a human and robot interacted over several minutes and measured the human's subjective experience of the event as well as how attentive the human was. These studies have made significant contributions to designing effective interaction protocols and determining design parameters for developing robots that can interact naturally with humans. However not all aspects of social attention can be captured with these methods. Overt behavior is only the end result of an entire cascade of sensory, cognitive, and decision making processes that occurs within the brain. For example human psychology research shows a complex array of processing occurs within a fractions of a second when we see another person shift their eye gaze or head turn [28, 29]. These findings suggest that we must study covert and unconscious mental processes that are involved in social attention in order to have a full understanding of human interaction.

To study the more automatic and covert aspects of social attention we used a variant of the well-established Posner spatial attention task [30] to compare robots' and humans' ability to direct a human observer's attention. This paradigm measures a human's ability to follow a directional cue using reaction time and is sensitive to events occurring in millisecond timescales. The Posner paradigm was used to study attentional systems and typically uses an arrow cue to direct attention to the periphery. Targets that appear in the cued location can be reported faster compared to targets appearing in an uncued location. This paradigm is simple but it gives us a window into automatic and often covert mental processes, such as our ability to direct attention based on social information. The Posner paradigm was adapted to study social attention through the use of cues such as eye gaze or body orientation. Our attention is directed to the same direction that another person is looking or turned [28, 29], though it has been suggested that social cues engage a more specific attentional system than cues such as [31, 32].

The Posner paradigm has been used in a few studies in the past to study human perception of robots [33, 34] and provides an objective measure (reaction time to reporting target location) that is useful as a gauge of how effective a robot is at manipulating automatic attention orienting processes that lie at the root of social attention. The present study extends past work by using dynamic video stimuli in addition to static images allowing us to examine the role of both robot form (physical appearance) and motion (motion kinematics). Form and motion are both features of a robot that may differ from humans and could change how the robot manipulates social attention. From the influential Uncanny Valley Hypothesis [5, 6], a near-human form is theorized to decrease the likeability of a robot. Traditionally, adding in unnatural motion kinematics aggravates the problem but more recent work on the effect of motion yielded differing results [7, 8,

22, 35]. It is possible that unnatural form or motion may affect a robot's ability to engage the automatic attention orienting processes. Manipulating the form and motion of artificial agents can allow us to determine which factor drives differing reactions to real humans *vs*. artificial agents.

To achieve the form and motion manipulation, we used well-controlled stimuli that were used in previous work [22], featuring three different agents, two robots and one human (Figure 1). For convenience of naming, we refer to the machine with a mechanical appearance as Robot while we call the machine with a biological appearance as Android. These stimuli are well controlled as the Robot and Android, actually the same machine, shared identical motion kinematics, and the Android and Human were highly similar in appearance. Interestingly, the study of social attention in general has not used moving stimuli, and therefore the present study provides an important bridge between vision science and interaction studies.

We may observe the following possible outcomes: an appearance driven effect where the Human and Android pair together in behavioral responses, a motion driven effect where the agents with the Robot and Android pairing together, an agent identity effect where the Robot and Android again pair together due to the subject knowing that they are both machines, and a mismatch effect where the Human and Robot pair together due to having matching form-motion biologicalness while the Android stands out as it has a mismatch between its appearance and form. Excluding the agent identity effect pattern, all of the other three effect patterns can be attributed to either form or motion. We hypothesize that the appearance or motion of an agent may influence its ability to manipulate social attention. Our experiment recorded reaction time as the dependent measure and can measure multiple aspects of social attention.

1.1 Attention Cueing
We can measure the *attentional cueing* ability, or the effectiveness of each agent in directing spatial attention. This is an index derived from how much faster (or slower) subjects respond targets that appeared on the cued side (Valid Cue) of an agent compared to targets appearing on the opposite side (Invalid Cue). Such orienting occurs automatically, without any body movements, nor even require eye movements on the part of the subject (covert attention). The Posner paradigm is able to track covert shifts in attention as subjects respond faster to targets that appear on the Validly Cued side compared to the Invalidly Cued side as their attention is automatically directed to the valid side and thus they are prepared to respond to targets appearing there.

Spatial cueing is an aspect of social attention that has been studied in the past [36]. Most work to date has explored gaze cues. In fact, the mere presence of eyes may be sufficient to trigger attentional orienting [37]. There is a smaller literature on studies like the present one using head turn cues. Head and gaze cues may make separate influences on the orienting of attention, though both types of cues are powerful cues for social attention [38, 39]. More relevant to HRI, cueing of attention with agents that have non-human form has been studied but it is unclear whether increased biologicalness facilitates or impairs attentional cueing. In a gaze cueing study, schematic faces were found to cue spatial attention more than realistic faces [40], but this could be due to the eyes in the former stimuli providing a more clear and salient directional signal.

A small literature exists on attentional cueing with artificial agent stimuli. In Admoni and colleagues' work [33], robots were not found to cue human attention. On the other hand Chaminade and Okka [34], who used a similar paradigm, found that both robots and humans could cue attention. Furthermore, social attention studies using live viewing have also found that humans can follow a robot's attentional cues [16, 27].

1.2 Attention Capture
We can also observe *attention capture* effects of each agent by measuring overall reaction time to each agent. This measure can give us a relative estimate of how effectively the agent cue held onto the subjects' attention. This is because subjects had to process the agent and its cue first before being able to shift attention away from it and respond to the target. If the subject spends more time processing cues from one agent type (both Valid and Invalid) then it can be said that particular agent type captured attention more.

Attention capture is also an important aspect of social interaction. We want to direct attention at certain times during interaction but we also want to attract attention towards ourselves. If the appearance or motion of an agent influences attention capture then the design of the agent should be considered if the agent must be able to capture attention as part of its function. Furthermore, there may be cases, such as emergency response, where it is undesirable for a robot to capture too much attention. Therefore, investigating the relation between appearance, motion, and attention capture can improve our understanding of robot design.

1.3 Presence of Motion
In addition to the two experimental factors, we also manipulated the *presence of motion* in the stimuli (static vs. dynamic) across experiments, which allows us to explore the effect of motion cues on social attention. The effect of motion on attentional cueing in the Posner paradigm has not been well explored. Previous work on spatial cueing had used static stimuli. There is some work on the role of dynamic cues in the developmental literature, but they use very different paradigms than the present study. Studies with infants have suggested that motion plays an important role in helping infants acquire attentional orienting skills [41, 42] and dynamic cues may be necessary for cueing infants' social spatial attention [43].

2. Methods
2.1 Subjects
Subjects were recruited from the student body at the University of California, San Diego and had an average age of 21.5 years. Twenty one subjects participated in the Image Experiment (13 female). Twenty subjects participated in the Video Experiment (16 female). No subject participated in both of the experiments. Subjects gave written informed consent in accordance with the institutional review board of this UCSD prior to participating in the study.

2.2 Stimuli

The stimuli were images and videos of three agents (Figure 1); two artificial and one human [21, 22]. The three agents varied in their form and motion kinematics. This stimuli set has actions performed by a realistic robot, Repliee Q2, and the human actor that Repliee Q2's appearance was modeled on. Repliee Q2's external human skin was also removed for a subset of the videos resulting in a mechanical form with the exact same motion kinematics as the realistic form version of Repliee Q2 (since the 2 are the same machine). For convenience of naming, the Robot refers to Repliee with mechanical form and motion while Android refers to Repliee with human-like form and mechanical movement. Finally, Human refers to the human actor, which has both human-like appearance and movement.

Videos of the three agents were recorded with the same background, lighting, and camera settings. The videos were frontal view and restricted to the upper body of the agents. In this study the directional cue was a video of the agents performing a 45 degree turn. The action duration was two seconds for all agent types. Static image stimuli were generated by taking one video frame of the looking forward phase and one frame from the turning phase. All actions were originally taped as a movement to the right of the agent (leftward of the viewer). The images and videos were flipped horizontally to produce cues in both directions. All stimuli (images and video) were converted to grayscale and matched for intensity.

2.3 Experiment and Procedure

The two experiments used similar procedures. Experiments were run using a 19" Dell Trinitron monitor with a screen resolution of 1024x768 and a refresh rate of 90Hz. Subjects were seated with their eyes 30" from the screen. The agents subtended 3x5.6 degrees of visual angle, the target letter was the letter W and it appeared 6.4 degrees from the central fixation point of the screen.

Two experiments were run using a repeated measure design with Cue Validity, and Agent type as experimental factors. The first experiment (Image Experiment) used static image stimuli while the second experiment (Video Experiment) used dynamic video stimuli. Thus, there was also a between subjects "presence of motion" manipulation in addition to the within subject experimental factors. Subjects were able to practice the experiment and familiarize themselves with the stimuli. In both experiments, subjects were informed that both the Robot and Android were machines while the Human was a real person. In a single trial (Figure 2) the subject fixated on a central fixation cross and then observed a static image of the agent looking forwards for approximately 1 second. This is followed by a 100 ms gray screen and then the appearance of the cue stimulus, which could be an image or a video. After a variable time delay (SOA), of 200ms or 400ms or 600ms from the appearance of the cue stimulus, the target would appear to the left or right of the agent. In video experiments, the target appeared as the video was still playing, not after the video concluded. Subjects were allowed to move their eyes once the cueing stimulus appeared. Subjects made a speeded response indicating the location of the target letter, pressing the right arrow key if the target appeared on the right of the agent or the left arrow key if the target appeared on the left of the agent. The target appeared with equal probability on the cued (Valid Cue) or uncued (Invalid Cue) sides and subjects
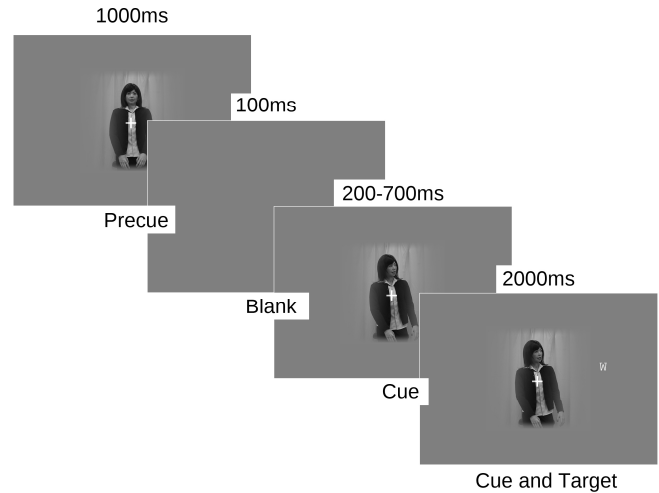
Figure 2: Timeline of a single experimental trial. The SOA during the cue period could be 200, 400, 600, or 700ms depending on the experiment. The trial terminated as soon as the subject responded to the target during the Cue and Target period.

were informed that the cue did not predict the target location. In 12.5% of the trials, the target would not appear, subjects were instructed to not respond on such trials. The purpose of this manipulation was to vary the task so subjects remained alert and did not default to making the same response repeatedly due to fatigue. Reaction times were measured from target appearance to key press.

2.4 Data Analysis

Only trials in which the target letter appeared were used for analysis. In all experiments, as typically done in attention studies, trials in which the subject made an incorrect response in indicating the location of the target letter were excluded from analysis. The correct trials in which the reaction time was faster than 200ms or slower than 4 standard deviations above the mean reaction time for a particular subject were also excluded from analysis. These are commonly used procedures in behavioral studies, and are done to ensure that the data reflect trials where subjects were alert and correctly performing the experimental task (i.e., helping to exclude accidental button presses or trials in which subjects were unusually distracted or slow). No more than 5% of trials for each subject were excluded from analysis for such reasons.

A repeated measures ANOVA was used to determine the effect of each experimental factor (Cue Validity and Agent) on reaction time. All experiments had two Cue Validity levels: Valid and Invalid, and three Agent levels: Robot, Android, and Human.
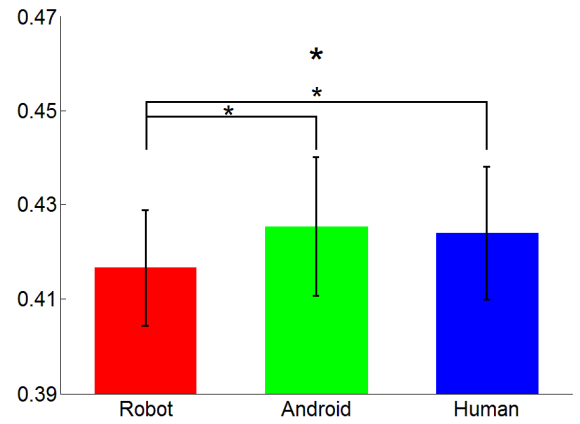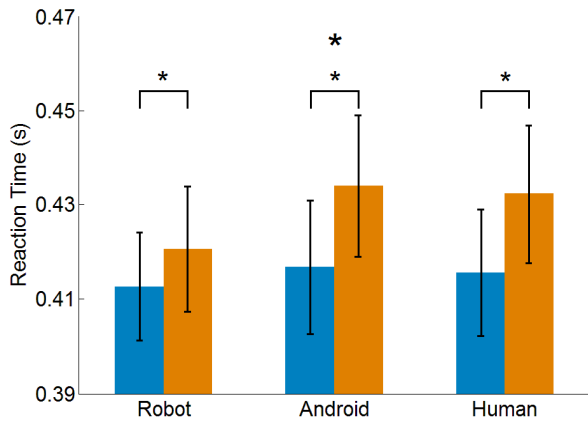
3. Results

In Image Experiment (Figure 3A) subjects were cued by static images of an agent making a turn towards the right or left of the screen. We found a main effect of Cue Validity $F(1,21) = 20.86$, p < 0.001. Pairwise t-tests find that Valid cues resulted in faster reaction times compared to Invalid cues (p < 0.001). Collapsing
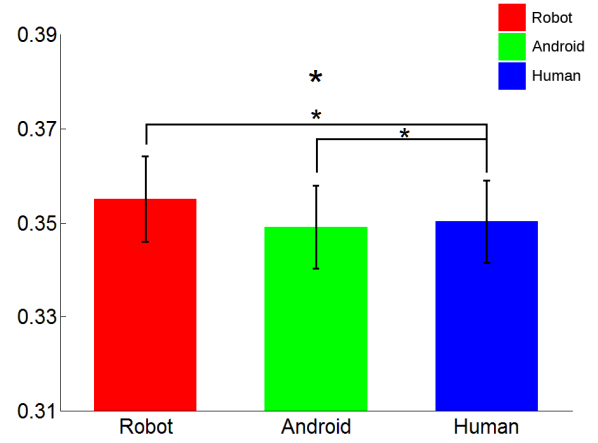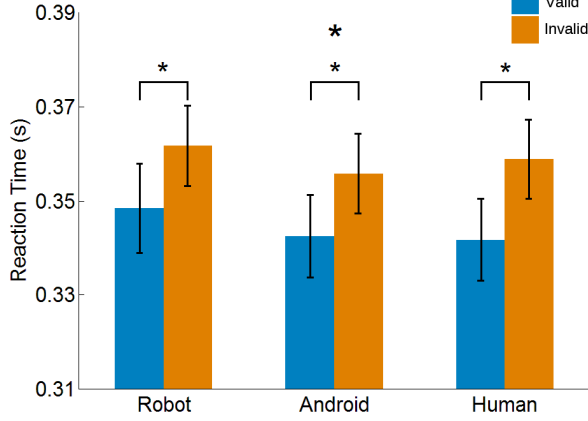
Figure 3: Results of Image Experiment (Figure 3A) and Video Experiment (Figure 3B). Column 1: Agent x Cue Validity Interaction. Column 2: Agent Main Effect. In both experiments, valid cues produced significantly faster reaction times that invalid cues though there was Agent x Cue Validity Interaction. Robot cues produced significantly slower reaction times compared to Android and Human cues in the Image Experiment, but the reverse was true for the Video Experiment.

across Cue Validity, we found a main effect of Agent $F_{(2,42)}$ = 7.76, p > 0.05. Pairwise t-tests between mean reaction times for each Agent find that the reaction times to Robot cues were significantly slower than the reaction times to Android and Human cues (Human vs. Robot p = 0.01, Android vs. Robot p = 0.002). Pairwise t-tests between reaction times to Android and Human cues were not significantly, p = 0.55. We found no Cue Validity x Agent interaction, $F_{(2,42)}$ = 1.145, p = 0.32.

In Video Experiment (Figure 3B), we found a main effect of Cue Validity $F_{(1, 19)}$ = 23.46, p < 0.001 with Valid cues producing faster reaction times than Invalid cues. There was also no Cue Validity x Agent Interaction $F_{(2,38)}$ = 2.08, p = 0.14. There was a main effect of Agent $F_{(2,38)}$ = 4.99, p > 0.05, though the effect was in the opposite direction compared to the Image Experiment. There was also no Cue Validity x Agent Interaction $F_{(2,38)}$ = 2.08, p = 0.14.

We also compared the Image and Video Experiment (Figure 4) in a between subjects model. This model tests for differences caused by the Cue Modality (whether the cue is presented as static or moving). We found a significant main effect of Cue Modality $F_{(1,40)}$ = 32.3, p < 0.001 with moving cues resulting in slower

reaction times that static cues. There was no Cue Modality x Cue Validity interaction suggesting that the cues were equally effective at cueing attention. There was a cue modality and agent interaction, $F_{(2,80)}$=3.24, p=.04, which further highlights the opposing agent main effects between Experiments 1 and 2.

4. Discussion
In the present study we investigated how the appearance and motion of agents influenced their ability to manipulate automatic, unconsciously directed social attention. We used a spatial attention task to explore both *spatial attentional cueing*, and *attention capture*. We did not find evidence that agent form or motion influenced the agents' ability to cue spatial attention. We did find overall *attention capture* effects driven by agent appearance. Furthermore, whether or not the agents moved changed which agent captured more attention.

4.1 Cueing Effect
As seen in the results, and discussed in more detail in section 4.3, whether the cue was static or dynamic made a difference in attentional capture and dynamics, but not in terms of spatial cueing. Among two previous experiments that used similar
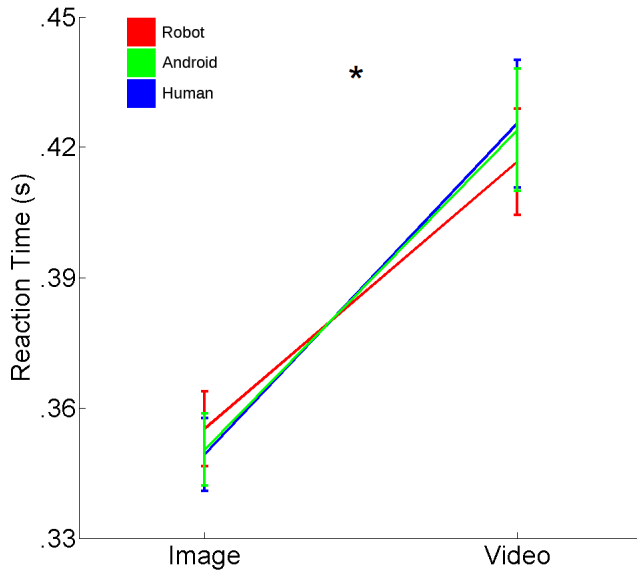
Figure 4: Agent x Experiment comparison, Reaction times were slower overall in the Video Experiment. Also visible is that reaction times to Robot Cues were longer relative to the other agents in the Image Experiment while the reverse was true in the Video Experiment.

paradigms to investigate attentional cueing with robots [33, 34], our findings were more consistent with [34] which found that robots could cue human attention as well as humans could. Our findings also complement studies using live interaction with a robot where humans could follow robots' attention cues during an interactive activity [16]. Though there are notable differences between the timescales explored with the Posner paradigm and live interaction experiments, both seem to suggest that robots can cue our spatial attention. As for the findings of [33], a couple possible explanations can be provided for the discrepancy. The cueing paradigm used in [33] had predictive target locations as well as four possible cueing directions rather than the typical two. Predictable targets may bias our attention [44, 45] and interact the robot's ability to direct attention. Predictability might be useful as a way of measuring how effective a robot is at creating or overcoming this bias relative to humans in future studies. In addition, the robot used in [33] was more toy-like and less mechanical in appearance. These factors may have impacted the end result of the study. Overall, it appears (at least in cases where the cue is not predictive of target location) the human spatial attention system does not "discriminate" against artificial agents, even those that do not look or move like humans.

4.2 Capture Effect

In addition to measuring attentional cueing, our experimental design also allowed us to measure attentional capture by the agent. Collapsing the spatial cues and looking at overall reaction times, we found an attention capture effect, meaning that the agents differed in how long they held onto attention (relative to each other). As discussed below in the next section, the presence of motion changed which agents captured more attention. But in both experiments the Robot agent captured attention differently form the Human and Android agents. This effect was based on the form of the agent since responses to the Robot differed from the other two agents, and the primary difference between the Robot

and the other two are in terms of the physical appearance (see Methods [21, 22]). Finally the overall longer reaction times in Video Experiment relative to Image Experiment suggests that moving stimuli in general capture attention for a longer period compared to similar static images. This is not surprising since motion requires additional processing and this can be reflected in the slowing down of reaction time. This falls in line with previous work that found that motion onset captures attention as objects that initiate motion are likely to be living and thus more important to the observer [46].

4.3 Presence of Motion Effect

In our study, we manipulated the *presence of motion* (static images vs. dynamic videos). The addition of motion in the cue drastically changed the attention capture effects we observed. There were two major effects, the first is that motion changes which agents captured more attention and the second is that, as mentioned above, reaction times to moving cues were longer than those to static cues (Image Experiment *vs.* Video Experiment). In both cases the Android and Human agent paired together in the results. These results contribute to the understanding of motion cues in spatial attention by revealing that attention operates on humanlike appearance or form in a similar fashion, despite a significant effect of motion.

Motion is integral to our studies as it is typical of real world HRI applications. A previous fMRI study found that observing an agent with realistic appearance triggered predictions of realistic motion; if the motion realism did not match the appearance, increased activity was observed in action perception networks of the brain suggesting that additional neural resources are activated to reconcile the prediction error [22]. Motion was also hypothesized to increase the effect of the Uncanny Valley [5, 6]. In the context of our study, if attentional mechanisms operate similarly to these prior studies, the reaction times to Android would differ relative to the Human with the addition of motion. Our results found a complex motion effect, but one that does not fit neatly with previous predictions on uncanny valley. It is therefore possible that social attentional mechanisms we probe here operate independently from the uncanny valley or similar phenomena. Furthermore, other researchers have also called into question the predictions laid out by Mori, and studies using ratings scales found that giving realistic agents motion did not always negatively impact their acceptability [7, 8, 35]. Our reaction time results also mirror some preliminary rating data of subjects viewing the same agents used here (Ürgen, Florendo, Saygin, unpublished). In these rating data, we found ratings of Human and Android humanlikeness and acceptability to become more similar with the addition of motion.

Our results suggest that the moving stimuli changed which features of the agent were prioritized in attention capture. One possibility is having two "sources" of salience that compete to control the attention capture effect when observing social agents. A low-level source based on visual features (such as spatial frequency, contrast), stimulus novelty, or a high-level source based on human-likeness. Stimuli with higher contrast can be more effective at capturing early attention [47] but on the other hand, stimuli with human forms are also more effective at capturing attention due to social relevance [48]. In our Image Experiment, the low-level salience source appears to "win" over the high-level source, and thus the Robot was more effective at capturing attention. In the Video Experiment, the fact that the

stimuli were moving may have boosted importance of high-level features. This may be due to the fact that in everyday life, moving objects with human appearance tend to be important. As a result, the moving agents with human form (Android and Human) could become more effective at capturing attention in the Video Experiment. In other words, the appearance and movement of the agent changes which features of the agent the brain predicts to be important and thus allocate attentional resources to.

We did not observe biological motion (the natural motion of the human agent) influencing attention capture or attentional cueing. It is unlikely that biological motion has no influence on attention as past work has found that attention and biological motion interact [32, 33]. It is possible the Human and Android agents were too similar in motion and the turning action did not provide sufficient biological motion cues. Further studies using more levels of human-likeness combined with subjective ratings can give insight into what exactly constitutes biological and non-biological motion and how they influence human social attention.

Though real social interaction is dynamic, to link the more naturalistic situations to those that have been studied in the literature, we applied the paradigm with both static and moving cues. Achieving this precise control of the agent's behavior is difficult in live interaction, especially when a human actor is involved. Although there are differences in experience between social interaction with a screen and with a live person or robot, results gained form this study can give us a better understanding about what kind of studies should be conducted live and what kinds can be done on screen. Screen based experiments are faster and cheaper to run and may prove useful for rapid prototyping of designs.

5. Conclusion

We investigated two aspects of automatic, unconscious social attention using both human a robot stimuli as social cues. These were *attentional cueing* and *attentional capture*. These are two important aspects of social attention as studied by experimental psychologists for decades, and index early, automatic, covert processes that are not possible to access with observational or rating studies. The current data shows that within 200ms of seeing an agent look in a direction, human attention is also turned to that direction. This early, automatic orienting suggests that even in situations where the human is not actively interacting with a robot, the attention systems of the brain are. This level of interaction is almost entirely unstudied in the HRI context. As such, our studies provide and important link between attention as studied in prior HRI studies with research in human psychology. More work will be required to explore other aspects of social attention and different methodologies are required as well to gain a complete understanding of how interaction with a human differs from interaction with a robot, and how to bridge the gaps. From the results of the present study, we speculate that the attention capture mechanisms in social attention are driven primarily by agent appearance, but motion can dramatically influence the exact features of an agent that take priority in capturing attention. One could suggest, if the goal is to create robots that elicit similar responses as humans elicit, then designing robots with biological appearances would be ideal; but this is one small piece of the whole of interpersonal interaction and additional parameters may change the design decisions, possibly varying also as a function of the robot's application environment.

The work presented here supports social attention research in both cognitive science and HRI. This work can also complement neural imaging studies to help link behavior with mental processes, and give insight into the neural mechanisms that operate when we encounter robots. Through these mutually supporting methodologies, we can build a holistic understanding of the mechanisms of human interaction, and eventually apply these to improvements in the design and development of 'neuroergonomic' social robots.

7. REFERENCES
[1] Coradeschi, S., Ishiguro, H., Asada, M., Shapiro, S. C., Thielscher, M., Breazeal, C., Mataric, M. J. and Ishida, H. Human-inspired robots. *Ieee Intelligent Systems*, 21, 4 (Jul-Aug 2006), 74-85.

[2] Breazeal, C. Toward sociable robots. *Robotics and Autonomous Systems*, 42, 3-4 (Mar 31 2003), 167-175.

[3] Kanda, T., Miyashita, T., Osada, T., Haikawa, Y. and Ishiguro, H. Analysis of humanoid appearances in human-robot interaction. *IEEE Transactions on Robotics*, 24, 3 2008), 725-735.

[4] Seyama, J. and Nagayama, R. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, 162007), 337-351.

[5] Mori, M., MacDorman, K. F. and Kageki, N. The uncanny valley [from the field]. *Robotics & Automation Magazine, IEEE*, 19, 2 2012), 98-100.

[6] Mori, M. The uncanny valley. *Energy*, 7, 4 1970), 33-35.

[7] Piwek, L., McKay, L. S. and Pollick, F. E. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, 130, 3 (Mar 2014), 271-277.

[8] White, G., McKay, L. and Pollick, F. Motion and the uncanny valley. *Journal of Vision*, 7, 9 2007), 477-477.

[9] MacDorman, K. F. and Ishiguro, H. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7, 3 2006), 297-337.

[10] Takano, E., Matsumoto, Y., Nakamura, Y., Ishiguro, H. and Sugamoto, K. The Psychological Effects of Attendance of an Android on Communication. *Spr Tra Adv Robot*, 542009), 221-228.

[11] Hanson, D., Olney, A., Prilliman, S., Mathews, E., Zielke, M., Hammons, D., Fernandez, R. and Stephanou, H. *Upending the uncanny valley*. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, City, 2005.

[12] Bartneck, C., Croft, E., Kulic, D. and Zoghbi, S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1, 1 2009), 71-81.

[13] Adair, J. G. and Spinner, B. Subjects Access to Cognitive-Processes - Demand Characteristics and Verbal Report. *J Theor Soc Behav*, 11, 1 1981), 31-52.

[14] Sardar, A., Joosse, M., Weiss, A. and Evers, V. *Don't stand so close to me: users' attitudinal and behavioral responses to personal space invasion by robots*. ACM, City, 2012.

[15] MacDorman, K. F., Minato, T., Shimada, M., Itakura, S., Cowley, S. and Ishiguro, H. *Assessing human likeness by eye contact in an android testbed*. City, 2005.

[16] Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H. and Hagita, N. *Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior*. ACM, City, 2009.

[17] Minato, T., Shimada, M., Itakura, S., Lee, K. and Ishiguro, H. Evaluating the human likeness of an android by comparing gaze behaviors elicited by the android and a person. *Advanced Robotics*, 20, 10 2006), 1147.

[18] Cheetham, M., Pavlovic, I., Jordan, N., Suter, P. and Jancke, L. Category Processing and the human likeness dimension of the Uncanny Valley Hypothesis: Eye-Tracking Data. *Frontiers in Psychology*, 42013).

[19] Seyama, J. i. and Nagayama, R. Probing the uncanny valley with the eye size aftereffect. *Presence*, 18, 5 2009), 321-339.

[20] Cheetham, M., Suter, P. and Jancke, L. The human likeness dimension of the "Uncanny Valley Hypothesis": Behavioral and functional MRI findings. *Front Hum Neurosci*, 52011), 126.

[21] Urgen, B. A., Plank, M., Ishiguro, H., Poizner, H. and Saygin, A. P. EEG theta and Mu oscillations during perception of human and robot actions. *Frontiers in neurorobotics*, 72013), 19.

[22] Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J. and Frith, C. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Soc Cogn Affect Neurosci*, 7, 4 (Apr 2012), 413-422.

[23] Tinwell, A., Grimshaw, M. and Williams, A. Uncanny behaviour in survival horror games. *Journal of Gaming & Virtual Worlds*, 2, 1 2010), 3-25.

[24] Birmingham, E. and Kingstone, A. Human social attention. *Prog Brain Res*, 1762009), 309-320.

[25] Wiese, E., Wykowska, A., Zwickel, J. and Müller, H. J. I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PloS one*, 7, 9 2012), e45391.

[26] Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H. and Hagita, N. *Footing in human-robot conversations: how robots might shape participant roles using gaze cues*. ACM, City, 2009.

[27] Mutlu, B., Forlizzi, J. and Hodgins, J. *A storytelling robot: Modeling and evaluation of human-like gaze behavior*. IEEE, City, 2006.

[28] Friesen, C. K. and Kingstone, A. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 51998), 490-495.

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

[30] Posner, M. I. Orienting of attention. *Quarterly Journal of Experimental Psychology*, 321980), 3–25.

[31] Hietanen, J. K. Social attention orienting integrates visual information from head and body orientation. *Psychol Res*, 66, 3 (Aug 2002), 174-179.

[32] Hietanen, J. K. Does your gaze direction and head orientation shift my visual attention? *Neuroreport*, 10, 16 (Nov 8 1999), 3443-3447.

[33] Admoni, H., Bank, C., Tan, J., Toneva, M., & Scassellati, B. Robot gaze does not reflexively cue human attention. . *Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Boston, MA, USA* 2011), 1983-1988.

[34] Chaminade, T. and Okka, M. M. Comparing the effect of humanoid and human face for the spatial orientation of attention. *Frontiers in neurorobotics*, 72013), 12.

[35] Thompson, J. C., Trafton, J. G. and McKnight, P. The perception of humanness from the movements of synthetic agents. *Perception*, 40, 6 2011), 695-704.

[36] Frischen, A., Bayliss, A. P. and Tipper, S. P. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol Bull*, 133, 4 (Jul 2007), 694-724.

[37] Quadflieg, S., Mason, M. F. and Macrae, C. N. The owl and the pussycat: gaze cues and visuospatial orienting. *Psychon Bull Rev*, 11, 5 (Oct 2004), 826-831.

[38] Langton, S. R. and Bruce, V. You must see the point: automatic processing of cues to the direction of social attention. *J Exp Psychol Hum Percept Perform*, 26, 2 (Apr 2000), 747-757.

[39] Langton, S. R. The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly journal of experimental psychology. A, Human experimental psychology*, 53, 3 (Aug 2000), 825-845.

[40] Hietanen, J. K. and Leppanen, J. M. Does facial expression affect attention orienting by gaze direction cues? *J Exp Psychol Hum Percept Perform*, 29, 6 (Dec 2003), 1228-1243.

[41] Farroni, T., Massaccesi, S., Pividori, D. and Johnson, M. H. Gaze following in newborns. *Infancy*, 5, 1 2004), 39-60.

[42] Moore, C., Angelopoulos, M. and Bennett, P. The role of movement in the development of joint visual attention. *Infant Behav Dev*, 20, 1 (Jan-Mar 1997), 83-92.

[43] Rohlfing, K. J., Longo, M. R. and Bertenthal, B. I. Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, 15, 3 (May 2012), 426-435.

[44] Naccache, L., Blandin, E. and Dehaene, S. Unconscious masked priming depends on temporal attention. *Psychological Science*, 13, 5 2002), 416-424.

[45] Wiese, E., Wykowska, A. and Müller, H. J. What we observe is biased by what other people tell us: Beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues. *PloS one*, 9, 4 2014), e94529.

[46] Abrams, R. A. and Christ, S. E. Motion onset captures attention: A rejoinder to Franconeri and Simons (2005). *Perception & psychophysics*, 68, 1 2006), 114-117.

[47] Parkhurst, D., Law, K. and Niebur, E. Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 1 (Jan 2002), 107-123.

[48] Langton, S. R. H., Law, A. S., Burton, A. M. and Schweinberger, S. R. Attention capture by faces. *Cognition*, 107, 1 (Apr 2008), 330-342.