Effects of Form and Motion on Social Attention

Maria Florendo[1], Alvin X. Li[1], Luke Miller[1], Ayse P. Saygin[1]

[1]Department of Cognitive Science, University of California San Diego, La Jolla, CA

INTRO

In this modern age, humans will soon encounter non-biological agents, such as robots, in their everyday lives. The purpose of integrating these non-biological agents into society is to assist with everyday tasks, such as teaching children in a classroom or helping patients in the hospital. If these robots are to have social roles within society, it is important that humans do not find them too distracting or unpleasant to work with. One example of social interaction that these robots could perform is directing spatial attention. Spatial orienting is directing another person's attention to a particular location in space using a visual cue, such as an arrow. Spatial attention is studied in the lab using the well-established Posner paradigm, in which subjects are asked to detect the location of a target letter. Reaction times are significantly faster if the cue direction and target location are the same compared to when the two are opposite (Posner & Peterson, 1980). Many studies have shown that spatial attention can be directed by social cues, such as a head turn or eye gaze (Driver et. al, 1999, Admoni, et. al, 2011). This is not surprising, seeing that orienting ones attention to a particular space using social stimuli is advantageous to social creatures because it provides essential information needed for survival. For example, if an animal sees another animal looking towards a certain direction, they can take this as a sign of possible danger.

The following social cues found in every day settings have both form and motion aspects. Surprisingly enough, previous studies that used the Posner paradigm have not utilized video cues as a way to present the cue to the subject. Instead, motion is implied through the presentation of two cues one after the other. Though previous studies have shown that non-biological agents can also cue attention, it has not been investigated whether the form and motion of these non-biological agents affect attention cueing (Chaminade, 2013). This is particularly interesting with the development of androids, or robots that have a human-like form, because it exhibits contrasting non-biological motion and biological form. Due to the conflicting relationship between form and motion, most people experience a feeling of eeriness or lack of empathy when perceiving androids.

The following study plans to investigate whether the form and motion of the cue affects its ability to direct spatial attention. Experiment 1, the study focuses on the influence of form on orienting attention by presenting biological and non-biological cues as static images like in previous experiments. Subjects are presented with still pictures of the agent looking forward followed by a picture of the agent turning its head to a certain direction. In Experiment 2, the study focuses on the interaction of form and motion and its influence on cueing ones attention by showing subjects videos of agents performing a head turn. The cue is performed by the same three agents with the following form and motion interactions: human (biological form and motion), android (biological form and non-biological motion), and robot (non-biological form and motion). In Experiments 3 and 4, the same paradigm is followed but the agents performed a non-directional bowing cue. Through the use of biological and non-biological stimuli, the following study highlights the effects of form and motion on spatial orienting.

LITERATURE REVIEW

Most studies that investigate directing of spatial attention use the following well established paradigm popularized by Posner, in which the task of the subject is to respond as quickly as possible to when they detect the target on the screen. Preceding the presentation of the target is a cue that either appears or directs the subject to a particular location on the screen. If the cued location is the same as the target location, then the subject detects the target more quickly and accurately than if the cued location is not the same as the target location (Posner, 1980).

Driver used this same paradigm and proved the reflexive nature of directing spatial attention. In the following experiments, they found the same results as Posner in that people were faster to direct their attention to a cued location despite having been given information beforehand that the location of the cue was uninformative or even misleading. In the first 2 experiments, subjects oriented faster to the cued side versus the uncued side even when they were told that the location of the cue was uninformative. In the third experiment, subjects still oriented faster to the cued location versus the uncued location despite knowing that the target would appear 4 times more likely on the opposite side as the cued location (Driver, 1999). These experiments show that orienting attention is reflexive.

Chaminade also used the Posner paradigm to investigate the effects of the form, or the social significance, of the agent on its ability to reflexively cue ones attention. The experiment used a human and an android, or a robot that looks like a human, as the cueing agents. Though their results showed an increased reaction time when the android agent was a cue compared to when the human agent was a cue, they did not find a significant interaction between agent type and cue validity. Therefore, the following study showed that both the robot and the android are able to direct spatial attention. They proposed that the increased reaction times for the android were simply because of attentional capture, in which the cue grabs the attention of the subject instead of directing it, leading to a slower reaction time for salient cues. This hypothesis is strengthened by the fact that they found similar results when they used human and android agents with their eyes blacked out (Chaminade, 2013). Therefore, Chaminade showed that the form of the agent can influence the cue's ability to orient attention.

However, for all of the previous experiments, the cues were all static in that motion was implied through the presentation of two images one after the other. For example, in Chaminade's experiment, the cues were presented as an image of the robot facing forward followed by an image of the robot's head turned to either the left or right. The current study looks to build upon Chaminade's experiments by investigating whether both form and motion of the cue influence its ability to direct spatial attention. It is important to study both of these aspects because cues that direct our attention in everyday social settings are dynamic. The study will use agents that demonstrate different interactions between form and motion, such as an android with non-biological motion but biological form. The theory by Masahiro Mori known as "The Uncanny Valley" hoped to explain the interaction between form and motion. This theory proposed that an increasing human likeness will result in increasing familiarity for people up until a point known as the uncanny valley (Mori, 1970). People experience a lack of empathy or a feeling of eeriness when perceiving agents that lie within this uncanny valley, such as androids. An explanation for this perception of uncanniness may be due to the interaction between form and motion, in which the human appearance of the android prompts people to expect natural movement. However, due to its non-biological motion, there is a failure to meet expectations (Saygin, et. al, 2012).

Given this, the following study examines the effects of form and motion on directing spatial attention. The study first attempts to replicate the findings of Chaminade in order to see if the form of the agent can influence cueing of attention. These studies are followed by the presentation of moving cues performing both directional and non-directional gestures to examine the effects of both form and motion in directing attention.

METHODOLOGY

Experiment 1

The first experiment investigates the effect of form on cueing attention through the presentation of static cues. In this experiment, subjects are asked to determine the location of a target letter W, either on the left or right side of the screen, by pressing the left or right arrow keys on the keyboard. Preceding the appearance of the target is an agent that cues the subject to either the left or right side of space. The schematic of each experiment trial is outlined in Figure 1. The subject is first presented with an image of the agent looking forward, or a precue image, for 1000 ms. This is followed by a gray screen with an ISI of 100 ms. Then, the cue image, or the image of the agent turning its head to the right or the left, appears on the screen followed by the target letter at varying SOAs. The presentation of two subsequent pictures is used to ensure that the study only investigated the influence of form on orienting attention. Subjects are informed at the beginning of the experiment that the location of the cue does not predict where the target will appear. Since reaction times are being measured, they are asked to respond as quickly and as accurately as possible.

Figure 1. *Schematic of Experiment Trial.* The trial first shows a precue, or the image of the agent looking straight at the subject, for 1000 ms. It is followed by a gray screen with ISI of 100 ms. Then the cue, or the image of the agent with its head turned, is shown. is shown and can either be an image of the agent with its head turned to the right, left, or still looking straight. Then after a variable SOA (200, 400, or 600 ms), the target letter W will appear either to the right or left of the cue.

The following study has a 3x3x2 experimental design with the following independent variables: cue validity, agent, and stimulus onset asynchrony. For each trial, the cue type is either valid, in which the cued location is the same as the target location, or invalid, in which the cued location is not the same as the target location. The agents are the following: robot (non-biological form), android (biological form), and human (biological form). Figure 2 shows the images of the three turned agent cues. The last independent variable is stimulus onset asynchrony, or the time from when the turning cue is initiated to when the target letter appears. The SOA is either at 200 ms, 400 ms, and 600 ms. SOA is studied because it is important to determine whether different agents could only cue spatial attention at certain SOAs.

Figure 2. *Images of turned agent cues*. The three agents used to cue the subject were a robot (non-biological form), android (biological form), and human (biological form).

RESULTS

The results were analyzed using a 2 sample T-test and 3-way ANOVA through R. Of the responses of the subjects, any responses that were incorrect, faster than 200 ms, or slower than 1500 ms were removed for analysis. There was a main effect found for cue validity ($F(1,21) = 20.68$, $p < 0.001$). Figure 3 shows that subjects responded overall quicker when the target and the cue direction were on the same side. Figure 4 further emphasizes this finding by breaking down each of these valid and invalid trials by agent type showing that all three of the social stimuli could direct spatial attention (Valid vs Invalid: $p<0.01$ for all three agents). However, Figure 4 also shows that there is no interaction between agent and cue validity ($F(2,42) = 1.14$, N.S). Thereby one cannot postulate that one particular agent can cue more effectively than others.

However, there was a main effect of agent found ($F(2,242) = 7.76$, $p = 0.002$). Figure 5 shows that collapsed across all trials, subjects responded overall slower to trials in which the agent performing the cue was the robot compared to when the android and human were performing the cue. Figure 6 shows no interaction between SOA x agent which concludes that there was a global difference in how the subjects responded to the agents at various SOAs ($F(4,84) = 1.145$, N.S.). These results are also seen when looking at the differences in cue validity, with no 3 way interaction of SOA, agent, and cue validity, signifying that all agents can cue at all SOA ($F(4,84) = 0.75$, $p = 0.56$).

The following experiment shows that all three of the agents could direct spatial attention. Though no one agent could direct attention more effectively than others, the following main effects of agent show that the form of the agent could affect how we perceive the following agents.

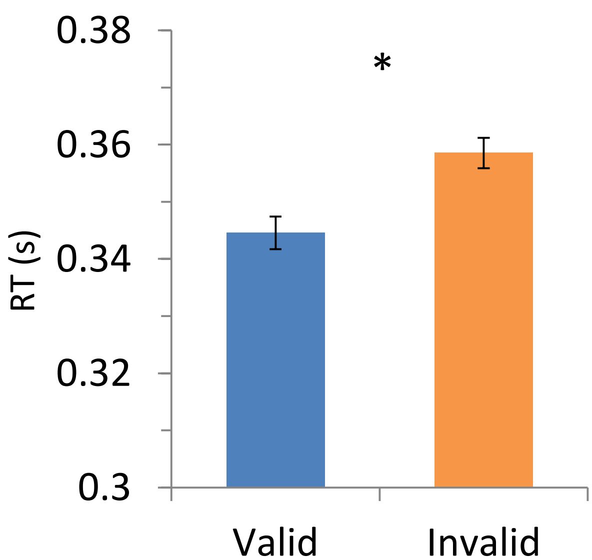


Figure 3. *Main Effect of Cue Validity for Static Cues.* Subjects responded significantly faster to valid cues, in which the target location and the cue direction were the same, compared to invalid cues, in which the target location and the cue direction were not the same $F(1,21) = 20.68$, $p < 0.001$.

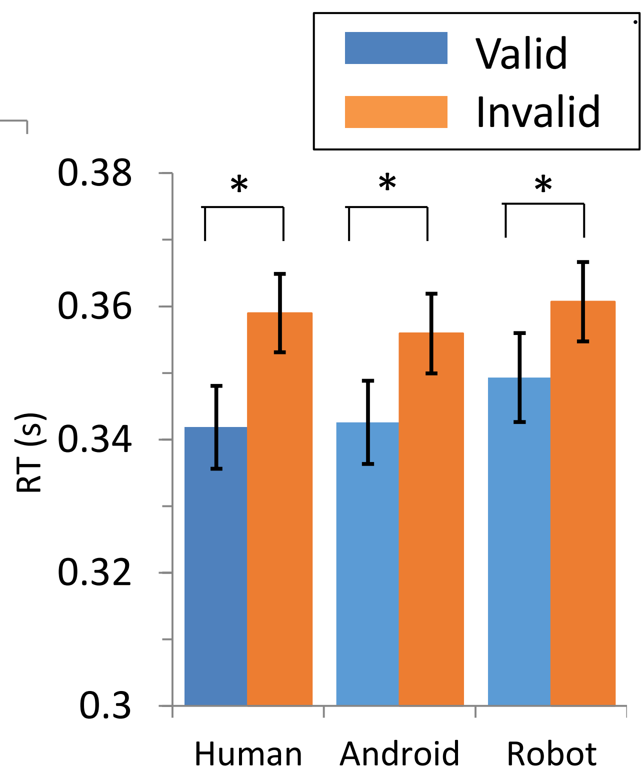


Figure 4. *No interaction between agent and cue validity.* Though for each agent subjects responded significantly faster for valid cues compared to invalid cues (Valid vs Invalid: $p < 0.01$ for all three agents) there was no interaction between agent and cue validity $F(2,42) = 1.14$, N.S. This shows that no one agent cued more effectively than the others.

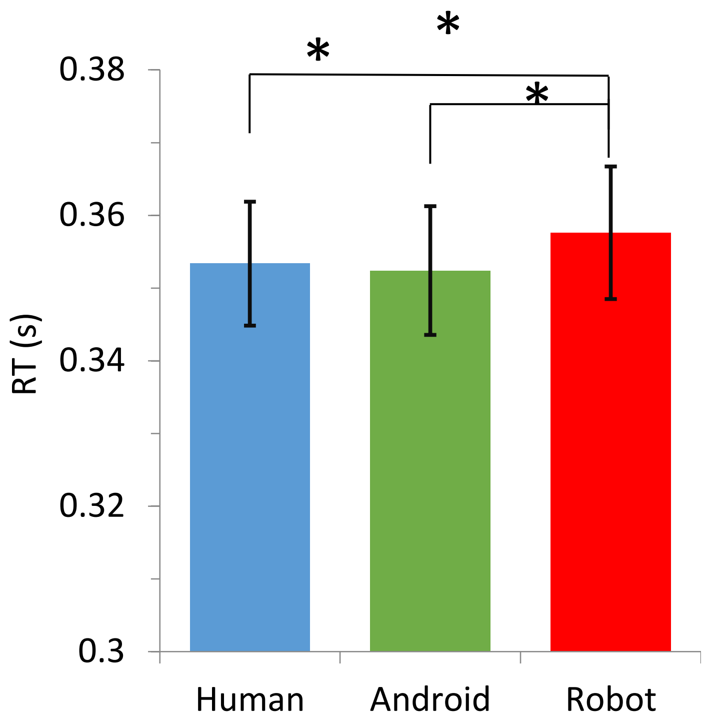
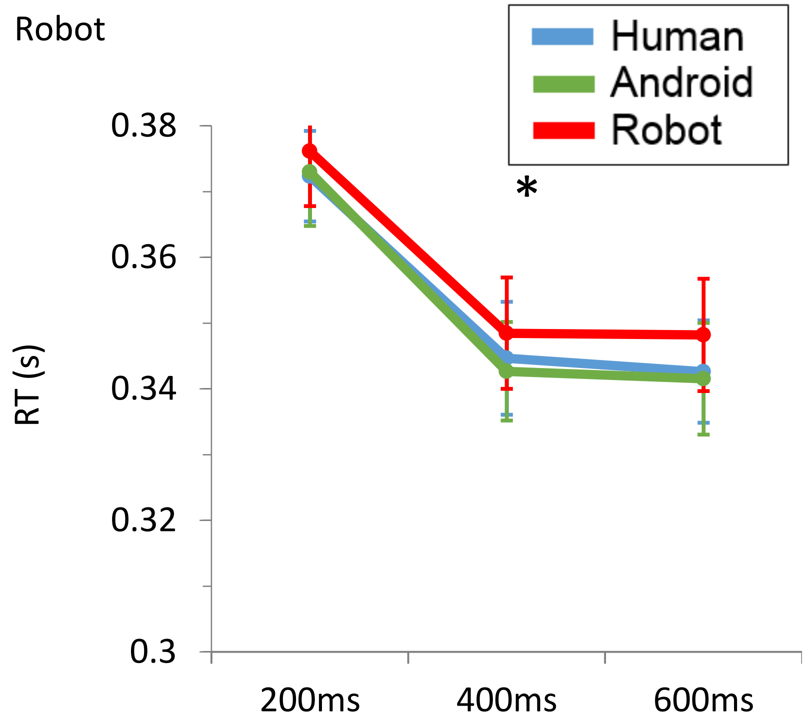


Figure 5. *Main effect of agent.* Collapsed across all trials, subjects responded significantly slower to robots compared to human and android F(2,242) = 7.76, p =0.002

Figure 6. *No interaction between Agent x SOA for static cues*. Subjects responded equally for all three agents at the different SOAs. (F4,84) = 1.145, N.S.



EXPERIMENT 2:

The second experiment investigates the effects of both form and motion of the cue on reflexive orienting. Subjects are presented with a video of three different agents turning their head either to the left or the right. After the video is presented, a target letter T appears either on the left or right side of the screen. The task of the subject is to determine the location of the target letter by pressing the left arrow button to indicate it appears on the left side of the screen or pressing the right arrow button to indicate it appeared on the right side of the screen. The experiment has a 3x2x2 design of agent x SOA x cue validity. The agents that cued the location are the following: robot (non-biological form and motion), android (biological form and non-biological motion), and human (biological form and motion). The two SOAs are 200 ms and 700 ms. Cue could be a valid cue, in which the cue direction is the same as target location in the same, or an invalid cue, in which the cue direction and target location is opposite. Subjects are once again told that the direction in which the cue turned was not predicative of where the target will appear.

RESULTS

The results were analyzed using a within subject ANOVA ran through R statistics. Once again there was a main effect of cue validity ($F(1,19) = 23.46$, $p < 0.001$) which showed that overall subjects responded quicker when the trials were valid versus invalid as seen in Figure 7. Figure 8 shows that all three of the agents could cue attention because there is a significant difference between valid and invalid trials within the agents (Valid vs Invalid, $p<0.01$ for all three agents). However, there was no interaction between cue validity and agent ($F(2,38) = 2.083$, p >0.05) showing that no agent could cue more effectively than the others. There was also a main effect of agent ($F(2,38) = 5.0$, $p = 0.01$) as seen in Figure 9. Collapsed across trials, subjects responded overall slower when the agent cueing their attention was the human and android compared to the robot. Finally, Figure 10 shows an interaction between agent x SOA in which at a shorter exposure of the cue, or at 200 ms SOA, there was an overall faster RT for when the agent cueing direction was the robot compared to the android and human ($F(2,38) = 9.28$, $p<0.001$).

The following results show that all three agents could direct attention as dynamically moving stimuli. However no particular agent could cue more effectively than the others. Figure 11 shows a comparison of experiments 1 and 2 in which the cues differed in their presentation (either static or moving). The following analysis shows an interaction between agent x experiment ($F(2,80) = 10.75$, $p <0.001$) such that for static cues, as seen in experiment 1, subjects had an overall slower RT to the cue performed by the robot and for moving cues, as seen in experiment 2, subjects had an overall slower RT to the cue performed by the android and human.

EXPERIMENTS 3 and 4:

The following experiment looks to investigate whether overall reaction time differences would occur irrespective of the spatial task. The previous paradigms are followed with experiment 3 by presenting cues in the same minimally dynamic fashion as experiment 1 and experiment 4 by presenting cues as dynamic videos. Instead of a directional head turning cue, a non-directional gesture of bowing is performed. Images are presented 50% mirror imaged to allow for analysis of side preference cued by the hand used for bowing.

RESULTS:

There was no main effect of agent as seen in Figure 12 for Experiment 3 ($F(2,38) = 2.15$) and for Experiment 4 ($F(2,38) = 0.02$). Therefore the reaction time differences for the different agents performing the cue are dependent on the context of a spatial orienting task.

DISCUSSION

The first experiment replicates the findings of Chaminade in which subjects responded significantly faster to a target that was at the same location as the cued side for all of the agents, including those that have non-biological form (Chaminade, 2013). Therefore, the first experiment shows that all agents, including robots and androids, could direct attention. However, the lack of interaction between agent and cue validity signifies that no agent could cue more effectively than others. It is also shown that all three agents are perceived similarly at different exposures to the cue. However, the fact that subjects overall responded slower when the agent presenting the cue was a robot compared to the android and human might show that the robot's salient form may be capturing attention instead of directing it. Due to its unfamiliar non-biological form, subjects may have focused their attention on the robot instead of directing their attention to the side that the agent was cueing. The findings from experiment 1 show that all three of the agents could direct attention through static cues and that form may play a role in directing attention by possible capturing attention. The following experiment further investigates the effects of these agents by looking at the influences of both form and motion in the agents' ability to direct attention.

Once again, the experiment 2 shows that all three agents could direct attention when presented in a dynamic fashion, or through the presentation of the video. However, due to the lack of interaction between cue validity and agent, it cannot be postulated that one agent can cue more effectively than the others. Despite this, overall reaction time differences are found for the android and human compared to the robot in that subjects responded overall slower when these two agents were performing the cue. It is hypothesized that this is caused once again by attentional capture, in that when subjects know that the agents will be moving, the agents with biological form may have more social relevance compared to the robot and therefore may result in the subjects having a harder time disengaging from these cues. The interaction between agent and SOA also shows the difference in dynamics of attention capture.

A comparison of experiments 1 and 2 shows that depending on how the cues are presented, subjects have a harder time disengaging from certain agents performing the cue. When the cues are presented in a minimally dynamic fashion, the robot's salient non-biological form may have captured the subjects' attention compared to the robot and android. When the cues are presented through moving video, the android and human's social relevance may have captured the subject's attention compared to the robot. With these results, it is interesting to wonder whether the following attentional capture occurs regardless of the spatial directing task. Experiments 3 and 4 investigate the following question by presenting a non-directional bowing cue. In experiments 3 and 4, all three of the agents captured attention equally. This shows that differences in disengaging from the agents are dependent on the context of cueing spatial attention.

In the following study, the effects of both form and motion on social cues of attention were investigated. It was found that all three of the agents, including non-biological agents, could cue attention. However, it is not exactly similar to their human counterpart because of differences in attentional capture. These results could have interdisciplinary applications by providing information to those developing non-biological agents to be used in social settings.

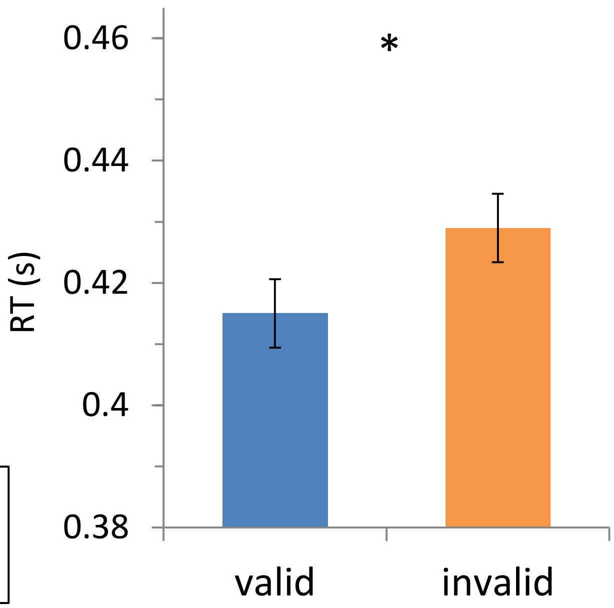


Figure 7. *Main Effect of Cue Validity for Moving Cues.* Subjects responded significantly faster to valid cues, in which the target location and the cue direction were the same, compared to invalid cues, in which the target location and the cue direction were not the same F(1,19) = 23.46 , p <0.001.

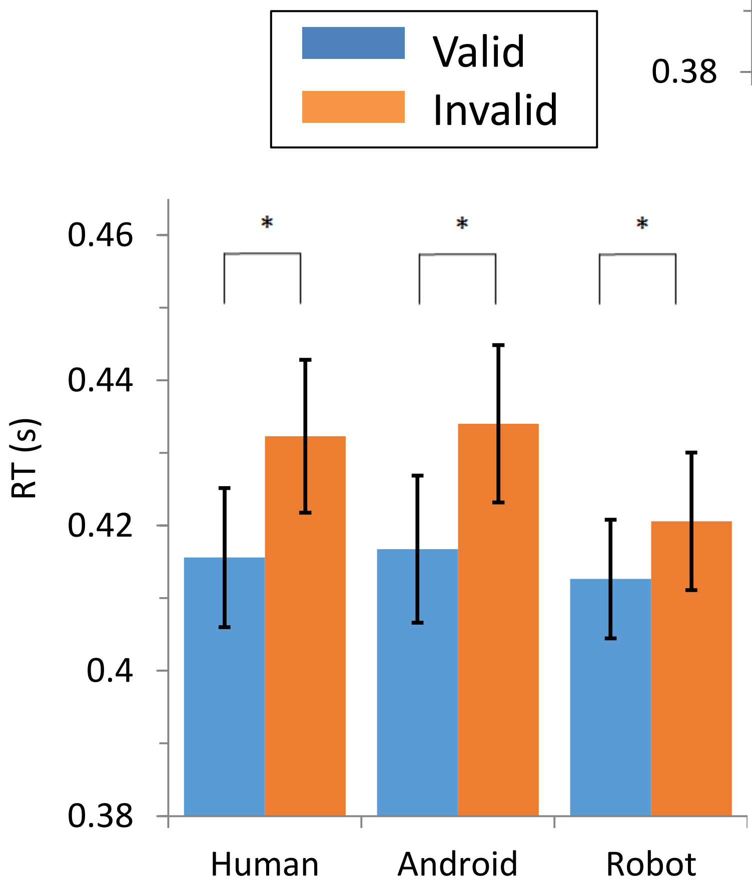


Figure 8. *No Interaction between cue validity and agent for Moving cues.* Despite subjects responding significantly faster for valid cues versus invalid cues for all three agents (Valid vs Invalid, p<0.01 for all three agents) there is no interaction between cue validity and agent (F(2,38) = 2.083, p >0.05.

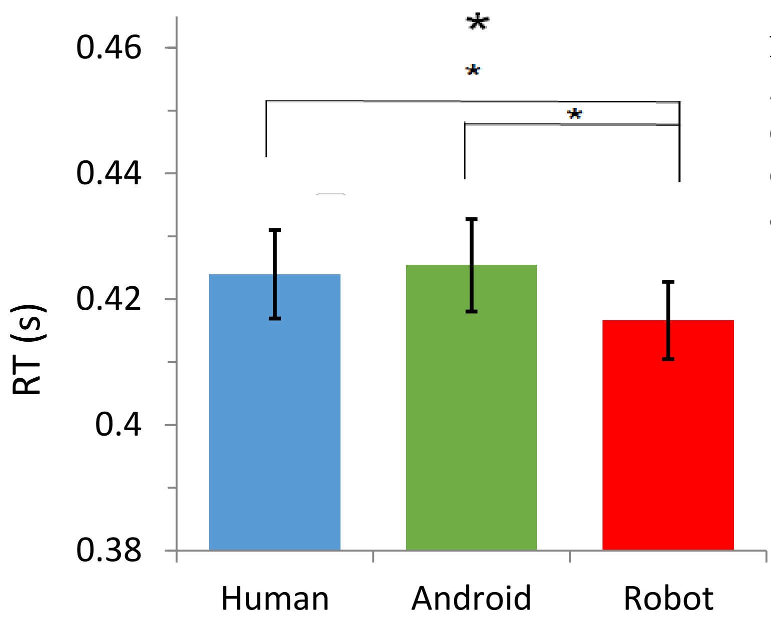


Figure 9. *Main effect for agent for moving cues.* Subjects responded overall slower to android and human cues compared to robot cues collapsed across all trials F(2,38) = 5.0, p = 0.01.

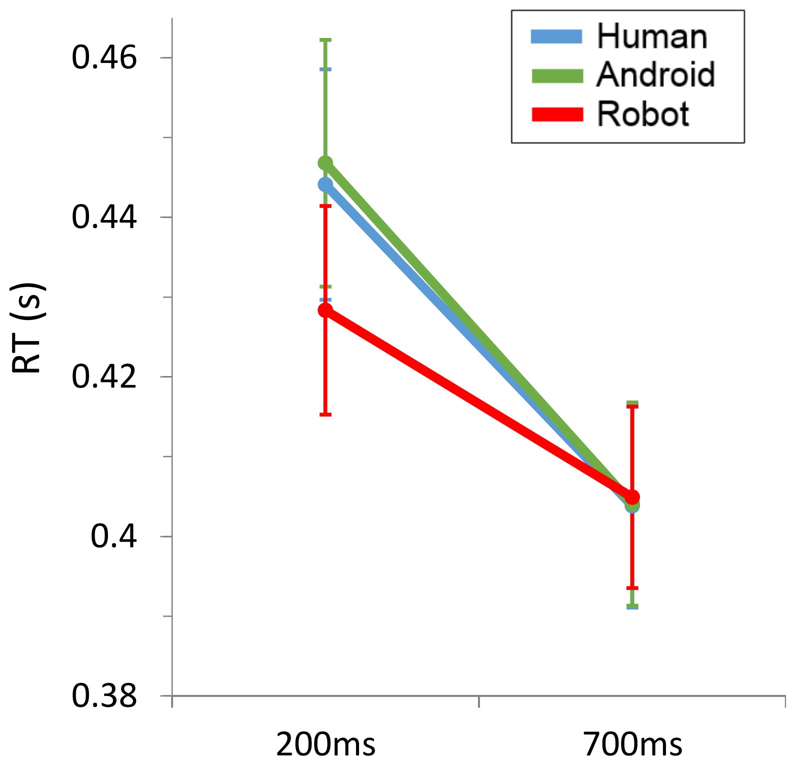
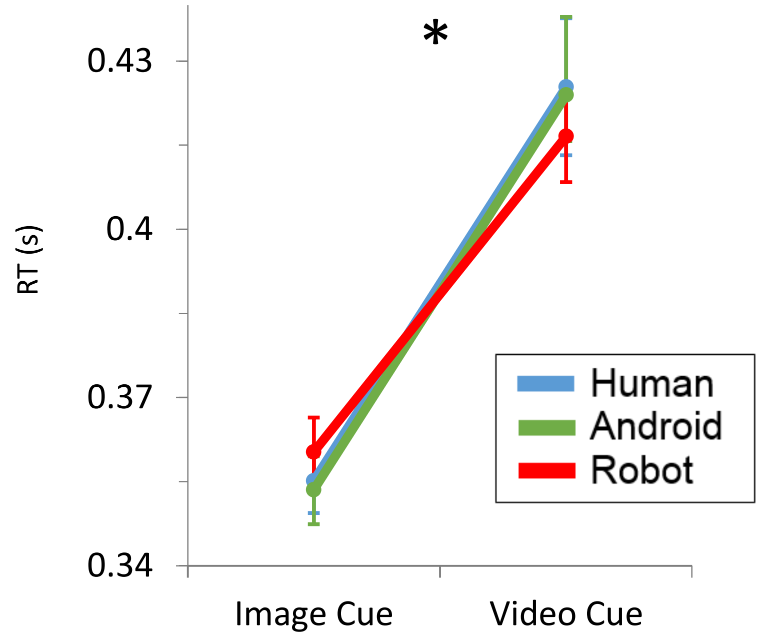


Figure 10. *Interaction between agent x SOA for moving cues.* At an SOA of 200 ms, subjects responded overall slower to human and android cues compared to the robot and responded equally to all three cues at the longer SOA of 700 ms (F(2,38) = 9.28, p<0.001.

Figure 11. *Interaction between agent x experiment.* Subjects responded overall slower when the cues were presented as videos compared to images. The overall reaction times for agents also differed between experiments, with robot having an overall slower reaction time when presented as image cues and human and android having an overall slower reaction time when presented as video cues $F(2,80) = 10.75$, $p < 0.001$.



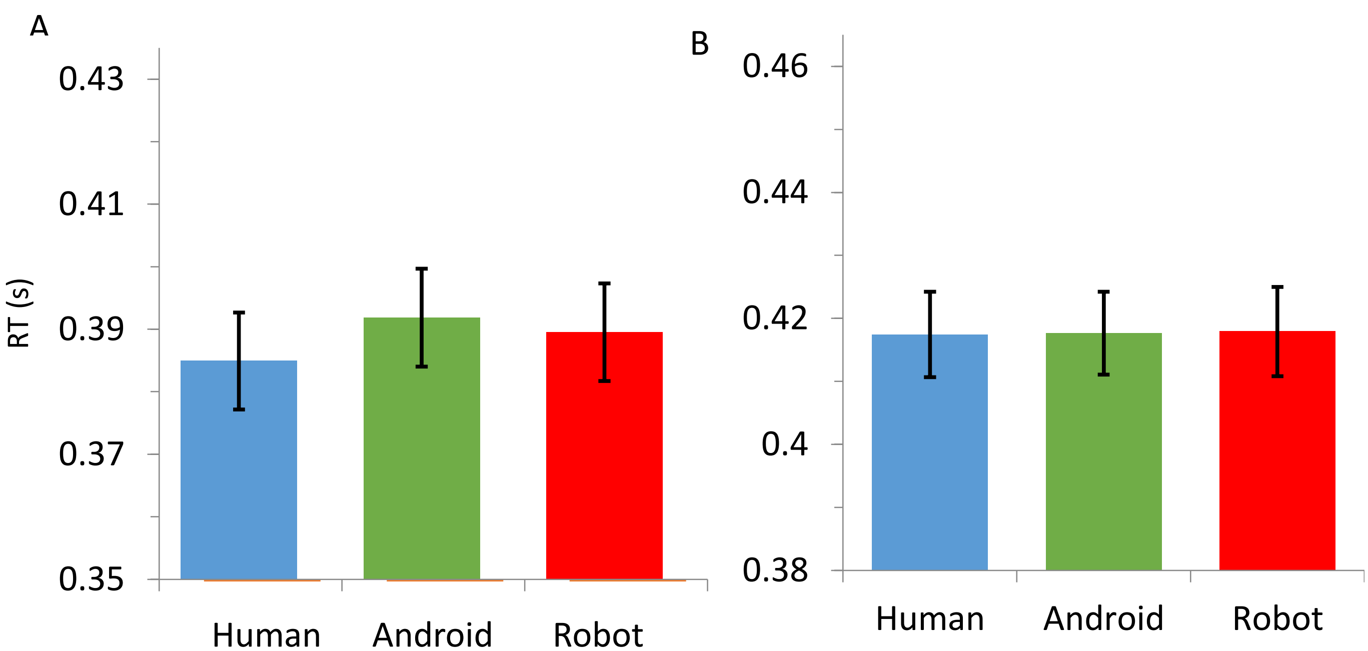


Figure 12. *No main effect of agent for both static and moving non-directional cues.*

When cues were presented as both static images and moving videos, subjects responded equally to all three f the agents when they were performing a non-directional bowing cue (A: Static images $F(2,38) = 2.15$) (B: Moving videos $F(2,38) = 0.02$).

WORKS CITED

Admoni,H.,Bank,C.,Tan,J.,Toneva, M.,and Scassellati,B.(2011).“Robot gaze does not reflexively cue human attention,” in Proceedings of the 33rd Annual Conference of the Cognitive Science Society (Boston, MA).

Chaminade, Thierry, and Maria M. Okka. "Comparing the Effect of Humanoid and Human Face for the Spatial Orientation of Attention." *Frontiers in Neurorobotics* 7.12 (2013): 1-7. Web.

Driver, Jon, Greg Davis, Paola Ricciardelli, Polly Kidd, Emma Maxwell, and Simon Baron-Cohen. "Gaze Perception Triggers Reflexive Visuospatial Orienting." Visual Cognition 6.7 (1999): 509-40. Web.

Mori, Masahiro. "The Uncanny Valley." *Energy* 7.4 (1970): 33-35. Web.

Posner, M.I. “Orienting of attention.” *Quarterly Journal of Experimental Psychology*, 32A (1980): 3-25. Web.

Saygin, A.P., Chaminade, T., Ishiguro, H., Driver, J. & Frith, C. (2012) The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. Social Cognitive Affective Neuroscience, 7(4):413-22.