How Does Varying Gaze Direction Affect Interaction between a Virtual Agent and Participant in an On-Line Communication Scenario?

Adam Qureshi[1], Christopher Peters[2], and Ian Apperly[3]

[1] Edge Hill University, Psychology Department Ormskirk, Lancashire, UK
[2] KTH Royal Institute of Technology, School of Computer Science and Communication, Valhallavagen 79, 100 44 Stockholm, Sweden
[3] University of Birmingham, School of Psychology Edgbaston, Birmingham, UK
qureshia@edgehill.ac.uk, chpeters@kth.se, i.a.apperly@bham.ac.uk

Abstract. Computer based perspective taking tasks in cognitive psychology often utilise static images and auditory instructions to assess online[1] communication. Results are then explained in terms of theory of mind (the ability to understand that other agents have different beliefs, desires and knowledge to oneself).The current study utilises a scenario in which participants were required to select objects in a grid after listening to instructions from an on-screen director. The director was positioned behind the grid from the participants' view. As objects in some slots were concealed from the view of the director, participants needed to take the perspective of the director into account in order to respond accurately. Results showed that participants reliably made errors, attributable to not using the information from the director's perspective efficiently, rather than not being able to take the director's perspective. However, the fact that the director was represented by a static sprite meant that even for a laboratory based experiment, the level of realism was low. This could have affected the level of participant engagement with the director and the task. This study, a collaboration between computer science and psychology, advances the static sprite model by incorporating head movement into a more realistic on-screen director with the aim of a.) Improving engagement and b.) investigating whether gaze direction affects accuracy and response times of object selection. Results suggest that gaze direction can influence the speed of accurate object selection, but only slightly and in certain situations; specifically those complex enough to warrant the participant paying additional attention to gaze direction and those that highlight perspective differences between themselves and the director. This in turn suggests that engagement with a virtual agent could be improved by taking these factors into account.

Keywords: Theory of mind, on-line communication, gaze direction, engagement.

1 Introduction

The aim of this study is to develop existing studies that have used static images in order to assess online communication [1] and have assessed gaze direction as a

[1] Online in this context means active interaction with an agent, not online as in internet-based communication.

R. Shumaker and S. Lackey (Eds.): VAMR 2014, Part I, LNCS 8525, pp. 305–316, 2014.

facilitatory factor [6], and relate the findings to engagement. This was achieved by manipulating the gaze direction of a director who instructed participants to select objects in an online communication scenario. This builds on [1] by incorporating head movement in order to change the gaze direction of the director and on [6] by including a condition that is predicted to inhibit correct object selection by the participant. As the rationale for the study and predictions derive from both theory of mind and communication literature, the introduction will cover each of these in turn, followed by a description of the prior studies methodologies.

1.1 Theory of Mind

Theory of mind is commonly defined as the ability to understand that others have different beliefs and mental states to you [11]. In addition, it also covers the ability to compute another person's perspective and to use the information from that perspective [12]. It also refers to being able to use that information to interpret behaviour or utterances.

The majority of psychological research on theory of mind has been developmental [12], and the classic test of theory of mind understanding is the false belief task [15]. Explicit understanding of false beliefs has been shown to develop by around 4-6 years old [15]. More recent studies have found that perspective-taking ability develops in infants [7] prior to the ability to pass false belief tasks.

The development of theory of mind has been associated with executive function and language ability [3] [5]. There is evidence that perspective-taking is cognitively efficient and automatic [12], but that more complex processes (such as false belief inferences) require executive processes [4]. Perhaps analogously, selecting perspectives is also thought to require executive processes [12]. This is supported by evidence from animal studies that show that chimpanzees exhibit low-level perspective taking abilities but are unable to understand false beliefs [2].

In summary, theory of mind may involve a cognitively efficient, perhaps innate, module that allows us to take perspectives. Dealing with different perspectives with respect to selecting which one to process and/or take information from may require executive processes – the latter develop over infancy, perhaps explaining why infants are able to pass simple (and implicit) tasks requiring perspective taking, but are unable to pass explicit false belief tasks until approximately ages 4- 6. An ability requiring executive resources is likely to be more flexible than one that does not, but due to this requirement it is also more error-prone and likely slower.

1.2 Communication

Theory of mind appears key to successful communication, as research on everyday communication [4] and conversational pragmatics [13] [14] suggest that to successfully communicate speakers and listeners must be able to take account of one another's knowledge, beliefs and intentions. However, this research assumes either implicitly or explicitly that such inference and use of information about mental states occurs quickly and efficiently. Psychological research on theory of mind suggests that inferring mental states may well occur quickly and efficiently (c.f. cognitively efficient perspective taking) but that using that information may not (c.f. executive resources needed for perspective selection).

During communication in real life, listeners can often see a speaker's eyes, and evidence suggests that information about eye gaze is used rapidly online to resolve ambiguous reference [6]. Results suggest that eye gaze has an automatic, reflexive orientating effect on attention. However, this effect may not always be reflexive or hard-wired. There is evidence that eye gaze is a flexible cue that can be rapidly remapped – it is a source of information whose use can be modified according to communicative context. Eye gaze may then help in using information from a perspective once that perspective is calculated.

1.3 Experimental Method

Experiments in cognitive psychology often use static imagery to assess communication. The task developed in [1], based on a study by [7] required participants to move objects in a 4 x 4 grid as instructed by a director with an different perspective to the participant (Fig.1).

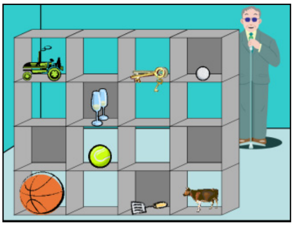

Fig. 1. Relational trial - instruction 'move the small ball one slot down'

The correct object was mutually visible to both the participant and director, whereas the competitor object was visible only to the participant. However, the competitor object was often a better referent to the instruction given.

Participants were shown the grid as it appeared from their own and from the director's perspective, so emphasizing the perspective-taking component of the task.

Critical instructions required the participant to make simple theory of mind inferences as they needed to calculate the perspective of the director and to use information from that perspective to interpret the instructions correctly.

Results showed participants made errors in object selection. The pattern of results suggested that they were able to calculate and take the director's perspective, but did not efficiently use the information gained. These errors suggest that participant fail to restrict the domain of potential reference to the 'common ground' of objects that were mutually visible / mutually known to themselves and the director.

Errors were attributed to egocentric bias: initial interpretation of an instruction (in this case) is based on information available to oneself, and "common ground" is optionally used to detect and correct errors from this initial interpretation [1] [7].

However, the director in the study was a static computer sprite wearing sunglasses. This experiment does not therefore include many cues which are normally part of conversation, such as eye gaze. As the eye gaze of a speaker is processed before the linguistic point of determination [6], this could suggest that participant's initial starting point is not egocentric: they attempt to use eye gaze to help resolve ambiguity before they hear any instructions. This means they do take information from other perspectives into account if a.) it is useful and b.) they are given the opportunity to do so. Processing of eye-gaze is low-level and relatively automatic (though can involve flexibility and remapping), but it is still a constraint whose influence can be weighted differently given the nature of the communicative context (c.f. other conversationally based sources of information).

This study [6] used instructions / referents that were ambiguous as to which object should be identified, but the eye gaze of the (human) director was always at the target object (facilitatory). The study that the current experiment is based on removed eye gaze as a factor by having the director wear sunglasses [1]. The current experiment builds on both studies by examining how participants deal with eye gaze as both a potentially facilitatory and inhibitory factor in object selection, and how this may be taken as a proxy for engagement with the director and their instructions.

2 Method

We used the design of [7] and [1] of a 4x4 grid, and systematically varied the gaze direction of the director in three conditions (the white square indicates the quadrant that the director is looking at, focusing on the center of that quadrant):

- Focus on mutually visible object (critical gaze, Figure 2 (R))
- Focus on an occluded object (competitor gaze, Figure 2 (L))
- No head movement (director looks straight ahead throughout the experiment)

Fig. 2. Relational trial - instruction: 'move the small ball down one slot'. (L): Competitor gaze. (R): Critical gaze.

Participants were shown an example grid from their own perspective and also from the director's perspective, and were explicitly told that the director would have no knowledge of objects that were in slots that were covered from their side and that they could not see. Therefore participants were made aware that there was a difference in perspective between themselves and the director, and that it might be necessary to take the director's perspective into account in order to respond correctly.

The instructions took the form of critical, control and filler questions. Critical and control questions were either ambiguous (referring to one of two items with the same name) or relational (referring to one of three items differing in spatial location or size). Control and critical grids were identical apart from a filler item replacing the critical item in the former (Figure 2 (L) shows a critical relational grid with a golf ball as the best referent for the instruction, but the tennis ball as the correct response when taking the director's perspective into account; Figure 2 (R) shows a control relational grid with no object present in the top-right slot (in the actual experiment this slot was occupied by an airplane). The best referent for the control instruction (move the small ball down one slot), from both the participant and director's perspective, is the tennis ball).

We investigated how participants process eye gaze by using three gaze direction conditions: (1) Critical: when the gaze of the director focused on the center of the quadrant containing the correct object (mutually visible to both participant and director). (2) Competitor: when it is focused on the quadrant containing the competitor object (only visible to the participant, so possibly increasing interference between self perspective and information). (3) When it has no informational value to the participant (default condition, as in the original experiment).

Based on the pilot study of [1], which showed that static cartoon eyes (with no head movement) of the director was enough to give the impression that he was referring to specific objects in the array, we assume that gaze direction is an additional source of information available when computing another persons perspective. The default condition in this experiment has the director looking directly at the participant. There should be no information from that gaze direction that can be used by the participant with respect to object selection.

Previous work used a starting point of the director looking at the center of the grid. Anecdotal evidence suggested that some participants did not look at the director or take into account the head movement. In this study each trial with the director looking at the participant in order to capture attention and promote engagement. The control condition was also changed from prior studies, where the director's gaze was focused on the centre of the 4x4 grid, to one where the director looked directly at the participant for the duration of the trials. Again, this was due to anecdotal evidence that participants could not distinguish where the director's gaze was directed (between the competitor, critical and grid-center control conditions).

We assume that the computation of the director's perspective occurs in conjunction with processing of gaze direction: it is additional information contained within the information gained by computing another's perspective. This information could then facilitate use of this information (by focusing on the critical object quadrant – accentuating common ground), inhibit use of this information (by focusing on the competitor object quadrant, possibly enhancing privileged ground of participant), or neither (focus on the participant).

If eye gaze is a constraint that can be weighted differently given the nature of the communicative context, then it would be expected that when the information from gaze direction does not disambiguate the reference to an object (critical gaze) or interfere with the selection of the correct object (competitor gaze), then the participant would not need to use the information from gaze direction but only the information gained from computing the perspective of the director. Therefore errors should arise from not using the information from the perspective of the director as in the original study (default condition).

When gaze direction contains relevant information for the disambiguation of the reference to an object or interferes with the selection of the correct object, we assume that the participant will use this information in addition to the perspective of the director. As this information cues participants to the quadrant in which either the critical or competitor object is located, we predict that if the information gained from computing the director's perspective is used, the critical gaze condition will increase accuracy and decrease response times in selection of the correct critical object (accentuate common ground; prediction 1). We also predict that the competitor gaze condition will decrease accuracy and increase response times in selection of the correct critical object (accentuate privileged ground; prediction 2). And finally, the default condition, where the gaze direction is focused on the participant, is not predicted to affect accuracy or response times (prediction 3). We also predict that the level of engagement between the participant and director could be indicated by increased speed and accuracy in the facilitatory gaze direction condition, and decreased speed and accuracy in the competitor gaze direction condition; therefore:

Prediction 1) Focus on mutually visible object (critical condition) = ↑ accuracy, ↓RT
Prediction 2) Focus on occluded object (competitor condition) = ↓ accuracy, ↑ RT
Prediction 3) Focus on participant (default condition) = ~ accuracy, ~ RT

3 Results

3.1 Error Rates

A 3 x 2 x 2 mixed ANOVA was conducted, with gaze condition (competitor v critical v default) the between-subjects factor and trial condition (control v experimental) and trial type (ambiguous v relational) the within-subjects factors.

There was a main effect of trial condition (control < experimental; $F (1, 56) = 215.41$, $p \leq 0.01$, $\eta p2 = 0.79$), a main effect of trial type (ambiguous < relational; $F (1, 56) = 92.57$, $p \leq 0.01$, $\eta p2 = 0.62$). There was no main effect of gaze condition ($F (2, 56) = 1.88$, $p = 0.16$, $\eta p2 = 0.06$).

There was no interaction between trial condition and gaze condition ($F (2, 56) = 1.45$, $p = 0.24$, $\eta p2 = 0.05$) or between trial type and gaze condition ($F (2, 56) = 1.51$, $p = 0.23$, $\eta p2 = 0.05$). There was an interaction between trial condition and trial type ($F (1, 56) = 73.33$, $p \leq 0.01$, $\eta p2 = 0.57$) but not between gaze condition, trial condition and trial type ($F (2, 56) = 1.38$, $p = 0.26$, $\eta p2 = 0.05$).

The interaction between trial condition and trial type was investigated by collapsing the gaze condition factor. Analyses showed that there was a significant difference between the control and experimental trial conditions in the ambiguous trial type condition (control < experimental; $F (1, 58) = 133.46$, $p \leq 0.01$, $\eta p2 = 0.70$) and in the relational trial type condition ($F (1, 58) = 171.80$, $p \leq 0.01$, $\eta p2 = 0.75$).

There was a significant difference between ambiguous and relational trial types (ambiguous < relational) in the control trial condition ($F (1, 58) = 10.44$, $p \leq 0.01$, $\eta p2 = 0.15$) and in the experimental trial condition ($F (1, 58) = 83.34$, $p \leq 0.01$, $\eta p2 = 0.59$).

The pattern of the interaction replicates that found in [1]. The error rates are shown in Table 1.

Table 1. Error rates by gaze condition, trial condition and trial type

	Control (mean (sd))		Experimental (mean (sd))	
Condition (n)	**Ambiguous**	**Relational**	**Ambiguous**	**Relational**
Competitor (20)	0.10 (0.31)	0.25 (0.44)	1.75 (1.41)	4.15 (2.43)
Critical (21)	0.05 (0.22)	0.29 (0.46)	2.19 (1.44)	4.81 (2.82)
Default (18)	0.11 (0.32)	0.33 (0.59)	2.11 (0.90)	5.78 (2.44)

3.2 Response Times

A 3 x 2 x 2 mixed ANOVA was conducted, with gaze condition (competitor v critical v default) the between-subjects factor and trial condition (control v experimental) and trial type (ambiguous v relational) the within-subjects factors.

There was no main effect of trial condition (F (1, 58) = 3.44, p = 0.07, ηp2 = 0.06), but there was a main effect of trial type (relational > ambiguous; F (1, 58) = 198.08, p ≤ 0.01, ηp2 = 0.77). There was also no main effect of gaze condition (F (2, 58) = 1.70, p = 0.19, ηp2 = 0.06).

There was no interaction between trial condition and gaze condition (F (2, 58) = 1.89, p = 0.16, ηp2 = 0.06) or between trial type and gaze condition (F (2, 58) = 0.27, p = 0.76, ηp2 = 0.01). There was an interaction between trial condition and trial type (F (1, 58) = 4.60, p = 0.04, ηp2 = 0.07) and between gaze condition, trial condition and trial type (F (2, 58) = 3.35, p = 0.04, ηp2 = 0.10).

The response times are shown in Table 2.

Table 2. Response time (ms) by gaze condition, trial condition and trial type

	Control (mean (sd))		Experimental (mean (sd))	
Condition (n)	**Ambiguous (ms)**	**Relational (ms)**	**Ambiguous (ms)**	**Relational (ms)**
Competitor (20)	2641.56 (451.77)	3422.32 (797.87)	2892.04 (685.24)	3198.75 (669.06)
Critical (21)	2545.97 (229.90)	3098.36 (288.70)	2800.05 (381.31)	3262.76 (515.25)
Default (20)	2566.35 (288.75)	3028.32 (326.88)	2582.67 (291.97)	3076.08 (405.73)

The interaction between trial condition and trial type was investigated by collapsing the gaze condition factor. Analyses showed that there was a significant difference between the control and experimental trial conditions in the ambiguous trial type condition (control < experimental; F (1, 60) = 15.70, p ≤ 0.01, ηp2 = 0.21) but not in the relational trial type condition (F (1, 60) = 0.00, p = 0.99, ηp2 = 0.00).

There was a significant difference between ambiguous and relational trial types (ambiguous < relational) in the control trial condition ($F_{(1, 60)} = 180.94$, $p \leq 0.01$, $\eta_p^2 = 0.75$) and in the experimental trial condition ($F_{(1, 60)} = 41.62$, $p \leq 0.01$, $\eta_p^2 = 0.41$).

This is shown in Figure 3:

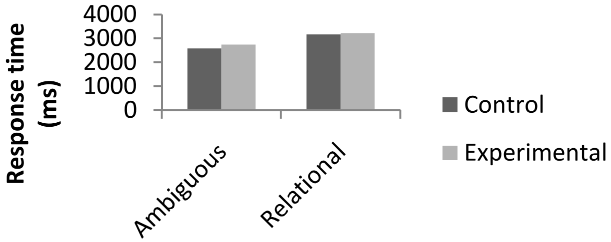

Fig. 3. Response time by trial condition and trial type (collapsed across gaze condition)

The interaction between gaze condition, trial condition and trial type was initially investigated by separate 2 x 2 within-subjects ANOVAs (trial condition x trial type).

Competitor Gaze Condition. There was no main effect of trial condition ($F_{(1, 19)} = 0.02$, $p = 0.90$, $\eta_p^2 = 0.00$), but there was a main effect of trial type (ambiguous < relational; $F_{(1, 19)} = 57.37$, $p \leq 0.01$, $\eta_p^2 = 0.75$) and an interaction between trial condition and trial type ($F_{(1, 60)} = 6.21$, $p \leq 0.02$, $\eta_p^2 = 0.25$). There was a main effect of trial condition (control < experimental) in the ambiguous trials ($F_{(1, 19)} = 8.76$, $p \leq 0.05$, $\eta_p^2 = 0.32$) but not in the relational trials ($F_{(1, 19)} = 1.54$, $p = 0.23$, $\eta_p^2 = 0.01$). There was a main effect of trial type (ambiguous < relational) in the control trials ($F_{(1, 19)} = 54.72$, $p \leq 0.01$, $\eta_p^2 = 0.74$) and in the experimental trials ($F_{(1, 19)} = 5.45$, $p \leq 0.05$, $\eta_p^2 = 0.22$), though the latter effect was much smaller.

Critical gaze Condition. There was a main effect of trial condition (control < experimental; $F_{(1, 20)} = 7.34$, $p \leq 0.05$, $\eta_p^2 = 0.27$), a main effect of trial type (ambiguous < relational) $F_{(1, 20)} = 57.87$, $p \leq 0.01$, $\eta_p^2 = 0.73$) but no interaction between trial condition and trial type ($F_{(1, 20)} = 0.48$, $p = 0.50$, $\eta_p^2 = 0.02$). There was a main effect of trial condition (control < experimental) in the ambiguous trials ($F_{(1, 19)} = 8.95$, $p \leq 0.05$, $\eta_p^2 = 0.31$) but not in the relational trials ($F_{(1, 19)} = 2.06$, $p = 0.17$, $\eta_p^2 = 0.09$). There was a main effect of trial type (ambiguous < relational) in the control trials ($F_{(1, 19)} = 121.37$, $p \leq 0.01$, $\eta_p^2 = 0.86$) and in the experimental trials ($F_{(1, 19)} = 14.46$, $p \leq 0.01$, $\eta_p^2 = 0.42$), though the latter effect was smaller.

Default Gaze Condition. There was no main effect of trial condition ($F_{(1, 19)} = 0.47$, $p = 0.50$, $\eta_p^2 = 0.02$). There was a main effect of trial type (ambiguous < relational; $F_{(1, 19)} = 107.17$, $p \leq 0.01$, $\eta_p^2 = 0.85$), but no interaction between trial condition and trial type ($F_{(1, 19)} = 0.11$, $p = 0.74$, $\eta_p^2 = 0.01$). There was a main effect of trial type (ambiguous < relational) in the control trials ($F_{(1, 19)} = 95.04$, $p \leq 0.01$, $\eta_p^2 = 0.83$) and in the experimental trials ($F_{(1, 19)} = 37.68$, $p \leq 0.01$, $\eta_p^2 = 0.67$), though the latter effect was slightly smaller. There was no main effect of trial condition in the ambiguous trials ($F_{(1, 19)} = 0.16$, $p = 0.69$, $\eta_p^2 = 0.01$) or in the relational trials ($F_{(1, 19)} = 0.32$, $p = 0.58$, $\eta_p^2 = 0.02$).

A one-way ANOVA was conducted to investigate the differences between the gaze conditions for the different trial condition and trial type combinations. There was no main effect for control ambiguous trials (F (2, 58) = 0.46, p = 0.64, ηp2 = 0.02), experimental ambiguous trials (F (2, 58) = 2.18, p = 0.12, ηp2 = 0.07) or experimental relational trials (F (2, 58) = 0.63, p = 0.54, ηp2 = 0.02). There was a marginally significant effect for control relational trials F (2, 58) = 3.26, p ≤ 0.05, ηp2 = 0.10, though post-hoc tests revealed there were no significant differences (all p's > 0.05). The overall pattern of results is shown in Figure 4.

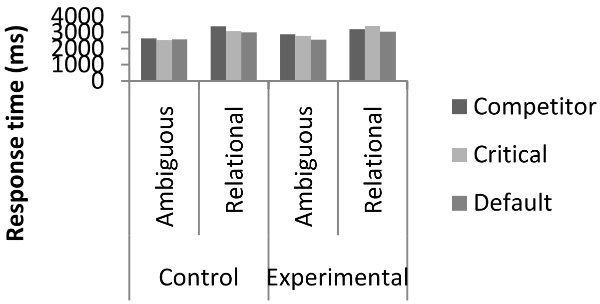

Fig. 4. Response time by gaze condition, trial condition and trial type

The error rates show no difference between the gaze conditions and replicate the findings of [1]. Although there are also no differences between gaze conditions in response times, there is an interaction between gaze condition, trial condition and trial type. Further analyses suggest that this interaction is caused by there being a significant difference between control and experimental conditions in the ambiguous trials for both the critical and competitor gaze conditions[2]. This is contrast to there being no difference between control and experimental conditions in the default gaze condition. The higher response times in experimental conditions would be expected due to the theory of mind inferences needed. However, this is not the case for relational trials in any gaze condition, though they do have a longer, but not significant, response time for both the critical and default gaze conditions. In the competitor gaze condition, the experimental relational trials are faster than the control relational trials, albeit not significantly so. This suggests that the director looking at the competitor object (in fact the wrong object that they have no knowledge of) helps the participant select the correct object, but only in the relational trial type (and in the experimental condition).

- Critical condition = no effect (no additional information after perspective calculation); experimental > control (amb
- Competitor condition = experimental relational is faster due to violation of expectation, ambiguous no difference
- Default condition = no effect (no information after perspective calculation).

[2] All gaze conditions show the same pattern of relational trials having a longer response time than ambiguous trials, both in the experimental and control trial conditions.

4 Discussion

The pattern of response times of the relational trials in the competitor gaze condition could be explained in terms of the violation-of-expectation paradigm that has been used with infants as an implicit measure: they will attend longer to situations that are unusual compared to those that they 'expect' or do not find unusual. Participants expect that if the director is going to look at an object that could be the referent of their instruction, it will be the one that they can see (critical gaze). If that is true, then the director looking at an object that is only visible to the participant would violate that expectation. This may make the perspective difference between the participant and the director more salient, and therefore lead to faster responses. However it is unclear as to why this only happens in the relational trials. The ambiguous trials follow the opposite pattern, where experimental trials are slower than control trials. It is possible that this is due to there being fewer items to process in this condition (2, as compared to 3 in the relational condition). If that is the case, it may be that participants only process any perspective difference through eye gaze at a certain threshold of complexity.

As the violation-of-expectation paradigm is often used in conjunction with eye-tracking equipment, it is possible that within the timeframe of the instruction, they may attend proportionally more to the gaze direction / head movement of the director in the competitor group than in the critical group. This would support the idea that they do not expect her to look at what is an occluded object from her perspective.

As it appears that participant only take eye gaze into account in certain situations, engagement may be promoted in online communication when a.) situations are sufficiently complex for participants to require the additional information it can provide, and b.) perspective differences between the participant and director are accentuated by the area or object the director is gazing at. However, it must be noted that these differences in response time are around 200ms and not significant, though the pattern is promising.

4.1 Future Work

Currently the director's gaze direction moves from looking at the participant to focus on the centre of the quadrant where the critical object is located (critical gaze), to focus on the centre of the quadrant where the competitor object is located (competitor gaze), or does not move and continues to be directed at the participant (default gaze). A further manipulation could involve the director focusing on any random quadrant that does not contain the critical object. Manipulation of eye gaze within head movements may also provide a more naturalistic representation of gaze direction, as static eyes within a moving head are not rated as realistic [9]. In order to assess whether participant are paying particular attention to head movement and gaze direction in the competitor gaze group (as suggested by a violation-of-expectation explanation) would require the use of eye-tracking equipment. This is currently ongoing work (a method without head movement has been created using Experiment Builder, the proprietary experimental design software for the EYELINK-1000 plus eye-tracking hardware.

The next step is to incorporate an agent with head movement into the experiment). A further plan is to replace the virtual agent with video capture of a human confederate who will act as the director (with head movement), and also to replace the virtual agent with an arrow that replicates the movement of the virtual agent with respect to prompting participants to attend to a particular location. The former will compare the level of engagement and attention participants have for a virtual versus a human agent, and the latter will compare the same between a virtual agent, human agent and a symbol that has inherent cuing properties. All of these will be done in conjunction with eye-tracking equipment.

5 Summary

Overall it appears that when gaze direction contains information that increases the saliency of any perspective difference, participants are faster in selecting the correct object. However, participants appear to either ignore or not use facilitatory information that could be gained from gaze direction, and this increase in engagement is shown only in the competitor gaze direction condition and in the relational trials. This could be explained by the violation-of-expectation paradigm, and a complexity threshold at which using information from eye gaze direction is triggered.

If processing of gaze direction and engagement was automatic, an effect of the facilitatory gaze condition on the speed and accuracy of object selection (either negative due to there being more information to process, or positive due to the information aiding the resolution of the instruction) would be expected. Therefore it is possible, with the current data, to suggest that processing gaze direction, and hence engaging further with the director, while fast and presumably then cognitively efficient, is a process that can be 'ignored' in certain situations. Specifically, it can be ignored where it provides no additional information to that gained by computing another's perspective or in situations where the level of complexity is not sufficient for participants to need the additional information it can provide. Aspects such as naturalistic blinking may increase engagement enough to 'override' any (conscious or unconscious) propensity to ignore gaze direction. This study is also of potential significance to human-machine interaction and engagement [16], for example, in collaborative situations between humans and artificial humanoid entities, where variations in their relative positioning with respect to each other and the environment, and accompanying perceived behavioural cues (real or artificial), may facilitate or hinder interaction performance.

References

1. Apperly, I.A., Carroll, D.J., Samson, D., Qureshi, A., Humphreys, G.W., Moffatt, G.: Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. Quarterly Journal of Experimental Psychology 63(6), 1201–1217 (2010)
2. Call, J., Tomasello, M.: Does the chimpanzee have a theory of mind? 30 years later. Trends in Cognitive Science 12(5), 187–192 (2008)

3. Carlson, S.M., Moses, L.J., Claxton, L.J.: Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. Journal of Experimental Child Psychology 87, 299–319 (2004)
4. Clark, H.H., Marshall, C.R.: Definite reference and mutual knowledge. In: Joshi, A.K., Webber, B., Sag, I. (eds.) Elements of Discourse Understanding, pp. 10–63. Cambridge University Press, Cambridge (1981)
5. de Villiers, J.G., Pyers, J.E.: Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief- understanding. Cognitive Development 17, 1037–1060 (2002)
6. Hanna, J.E., Brennan, S.E.: Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. Journal of Memory and Language 57, 596–615 (2007)
7. Keysar, B., Lin, S., Barr, D.J.: Limits on theory of mind use in adults. Cognition 89, 25–41 (2003)
8. Onishi, K.H., Baillargeon, R.: Do 15-Month-Old Infants Understand False Beliefs? Science 308(5719), 255–258 (2005)
9. Peters, C., Qureshi, A.: A Head Movement Propensity Model for Animating Gaze Shifts and Blinks of Virtual Characters. Computers and Graphics, Special Issue on Graphics for Serious Games 34(6), 677–687 (2010)
10. Peters, C., Qureshi, A., Apperly, I.A.: Effects of gaze direction of a virtual agent in an online communication game. European Society of Philosophy and Psychology, Ruhr-Universität Bochum (August 2010)
11. Premack, D.G., Woodruff, G.: Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences 1, 515–526 (1978)
12. Qureshi, A., Apperly, I.A., Samson, D.: Executive function is necessary for perspective-selection, not Level-1 visual perspective-calculation: Evidence from a dual-task study of adults. Cognition 117, 230–236 (2010)
13. Sperber, D., Wilson, D.: Relevance: Communication and Cognition, 2nd edn. Blackwell, Oxford (1986)
14. Sperber, D., Wilson, D.: Pragmatics, modularity and mindreading. Mind & Language. Special Issue on Pragmatics and Cognitive Science 17, 3–23 (2002)
15. Wellman, H.M., Cross, D., Watson, J.: Meta-analysis of theory-of-mind development: The truth about false belief. Child Development 72, 655–684 (2001)
16. Peters, C., Castellano, G., de Freitas, S.: An exploration of user engagement in HCI. In: Proceedings of the Affect-Aware Virtual Agents and Social Robots (AFFINE) Workshop, International Conference on Multimodal Interfaces and Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI 2009), Boston, MA, USA, November 6 (2009)