Comparing Behavioural and Self-Report Measures of Engagement with an Embodied Conversational Agent: A first report on Eye Tracking in Second Life

Sara Dalzel-Job*
Craig Nicol†
Jon Oberlander‡
Human Communication Research Centre, School of Informatics,
University of Edinburgh, 2 Buccleuch Pl., Edinburgh, EH8 9LW, UK

Keywords: eye tracking, embodied conversational agents, Second Life

1 Introduction

Embodied Conversational Agents (ECAs) are widely used to assist users in carrying out tasks. There are various reasons for including them in interfaces: they may help guide a user's attention to important information; they may improve overall task performance; or they may simply make the interface more attractive, thereby increasing the user's motivation to engage with a task.

An investigation into the eye movements of users during interaction with an ECA has shown that deictic gestures by an agent are an effective way of directing the user's focus of attention [Prendinger et al. 2007]. However, few studies into how people interact with Embodied Conversational Agents have examined eye movement data in conjunction with self-report and task-performance measures. In an unpublished exploratory study, it was found that users spend very little time attending to Genie, an ECA which assists users by guiding them along a route on a map; nonetheless, they reported that they preferred being guided by Genie rather than by voice alone [Dalzel-Job et al. 2006]. If the default is indeed for users to spend little time looking at an ECA, it is therefore of interest to discover what the effects might be of attempting to increase the amount of time users look at an agent while s/he is giving them instructions—by ensuring that he holds visual information that is critical to the successful completion of the task. We can then explore the relationship between looking at an agent, task performance and self-report measures.

However, there is a methodological issue which affects empirical studies of people interacting with ECAs. Much previous work has focused on systems that are not easily adaptable to domains other than those for which they were developed [Rickel and Johnson 1998].

Thus, while pursuing our research questions, we are also interested in developing a paradigm for investigating interaction with ECAs which supports a suitable level of adaptability. Second Life (SL) is a "3D online digital world imagined and created by its residents" [SL]. Using SL as a platform on which to investigate interaction with ECAs allows for a significant degree of adaptability to different domains, due to the versatile nature of the environment.

The current study therefore aims to extend the earlier results by developing an experimental paradigm that allows the recording and analysis of users' eye movements while they carry out tasks within Second Life. They are assisted in these tasks by Tracker, an avatar programmed to act as a robot. To the authors' knowledge, there have been no eye tracking studies within Second Life to date, and, as such, a key aim is to develop an experimental paradigm that enables the investigation of human users' eye movements during interaction with pre-programmed avatars (ECAs) and other users' avatars within SL.

2 Methodology

24 participants aged 18-40 (14 male) performed 3 tasks in 3 different rooms within a tower located on the Vue Island within SL [VUE]. Their eye movements were recorded using an Eyelink II (SR Research Inc.) eye tracker, recording at 500 Hz. Each participant was presented task-completion instructions by Tracker in 3 different forms: voice only, visible but providing redundant visual information (see *Figure 1*), visible and providing non-redundant visual information (see *Figure 2*). This produced 3 different *Tracker conditions*: *invisible*, *redundant* and *non-redundant*.

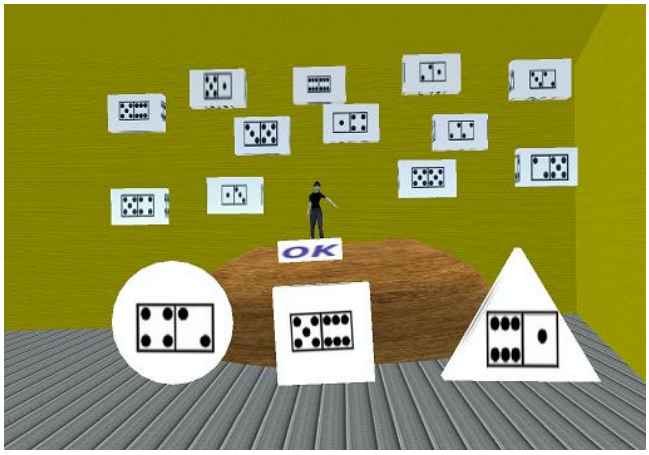

Figure 1: *"From the wall behind me, select a domino with the same total number of dots as the one on the triangle".*

The participants were instructed to select targets in the room that were in some way related to the stimulus to which Tracker was referring (either by pointing—non-redundant condition—or by name—redundant and invisible conditions). Subjects viewed the scene via *mouselook*, i.e. the view from their avatar's eyes, rather

*e-mail: S.Dalzel-Job@sms.ed.ac.uk
†e-mail: craig.nicol@bcs.org.uk
‡e-mail: J.Oberlander@ed.ac.uk

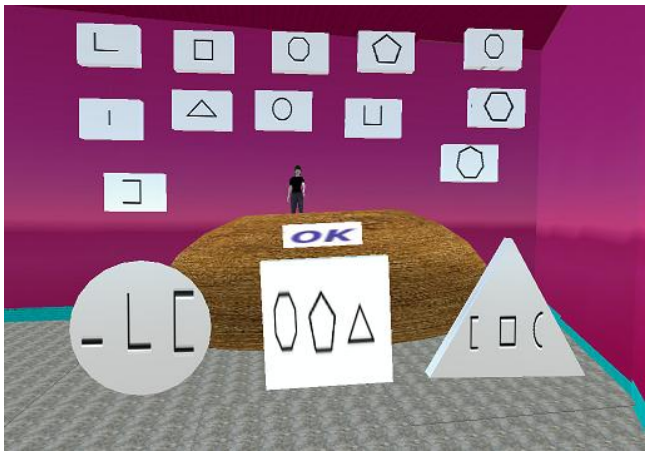

Figure 2: *"From the wall behind me, select a shape that you think best completes this sequence".*

than from beind its head. This ensured that their view was not occluded by the avatar, and that they were unable to move around the room, giving tighter controls over all participants. The dependent variables were: time spent looking at Tracker while he was giving instructions; number of tasks correctly completed; time taken to complete the tasks; and subjective ratings of the perceived helpfulness of Tracker for each task.

3 Results

Owing to the exploratory nature of this eye tracking study within SL, it emerged that there was a lack of dedicated software suitable for carrying out coding of the data. The data was annotated using NITE XML toolkit [Carletta et al. 2005]. Using this software, it was not possible to define a dynamic region of interest (ROI), meaning that it was generously approximated by the coder. This generous approximation was to maximise recall: to ensure that no looking at Tracker would be missed. Recall was important, because it was anticipated from the Genie study that little looking overall would occur. It must be emphasised that the current study was designed to help develop an experimental paradigm and, as such, any results can only be taken to be indicative of a within-study trend and cannot be reliably generalised to other studies.

Supporting the results of the Genie study, it emerged that participants looked very little at Tracker while he was talking, in both of the visible conditions. On average, less than 5% of this time was spent looking. There was significantly more looking at Tracker in the non-redundant than redundant condition. There was also a significant *negative* correlation between the amount of looking at Tracker and the number of correct targets selected in the redundant condition. This pattern was not found in the non-redundant condition. Participants reported Tracker to be more helpful in the non-redundant condition than in either the redundant or invisible conditions.

4 Conclusions

This study confirms that very little time is spent looking at an agent that is giving instructions, but that this small amount of looking can be increased by ensuring the agent holds task-critical visual information. It was found that, in the redundant condition, looking at an agent appears to be detrimental to task performance. This suggests that the question of whether an ECA should be included in an interface may depend on the type of task to be performed.

Second Life has proved to be an extremely versatile and useful platform for designing and running eye tracking studies. However, the current state of the available software is such that it is time-consuming and complicated to extract all of the data that could potentially be of great interest, with no guarantee that the results can be compared with any outside this study. Despite this, the current research has provided useful foundations for developing an experimental paradigm in SL that will allow researchers to evaluate systems that include, or consist of, ECAs assisting users in completing a task.

In comparing eye movements, task performance and self-report measures, it has been possible to uncover unexpected patterns of interaction between users and ECAs, as well as general views about the experimental design. This paradigm can now be revised to incorporate these findings.

References

CARLETTA, J., EVERT, S., HEID, U., AND KILGOUR, J. 2005. The nite xml toolkit: data model and query. *Language Resources and Evaluation Journal 39*, 4, 313–334.

DALZEL-JOB, S., PAPANANAGIOTOU, P., FOURNERON, N., VATTOLY, B., AND TAJWER, A., 2006. Examining the efficacy of "genie" the tour guide. Unpublished Edinburgh University MSc project.

PRENDINGER, H., MA, C., AND ISHIZUKA, M. 2007. Eye movements as indices for the utility of life-like interface agents: A pilot study. *Interacting with Computers 19*, 2 (Mar.), 281–292.

RICKEL, J., AND JOHNSON, W. 1998. Steve: A pedagogical agent for virtual reality. In *Proceedings of the Second International Conference on Autonomous Agents*, ACM Press.

http://secondlife.com/.

http://www.vue.ed.ac.uk.