Evaluating User Reaction to Character Agent Mediated Displays using Eye-tracking Technology

Mark Witkowski, Yasmine Arafa and Oscar de Bruijn

Intelligent and Interactive Systems Group
Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
Exhibition Road
London SW7 2BT
United Kingdom
{m.witkowski, y.arafa, o.debruijn}@ic.ac.uk

Abstract

Much work has been done to integrate visual character Agents into electronic display systems, notably in entertainment, CAE and e-retail contexts. In this study we investigated the usefulness of measuring users' eye movements as a tool for the evaluation of these character Agents. We particularly studied the usefulness of eye-tracking to obtain information about the capacity of the MAPPA character Agent to engage and direct the user's attention while providing both generalised system help and advice and when giving specific product information.

1 Introduction

This paper describes a pilot study undertaken to determine shifts of attention made by experimental participants while using the MAPPA Agent based system. Interaction with the user is mediated through an electronic Personal Service Agent (ePSA), an animated character Agent, which presents information to the user and embodies the personalisation of the system to each individual user. The MAPPA ("Multimedia Access through Personal Persistent Agents") system models an in-store, on-line, multimedia sales kiosk. MAPPA was initially developed to determine ways in which customer loyalty can be engendered and reinforced using the electronic sales medium (Arafa, *et al*., 2000, 2001; Pitt *et al*., 2000).

The MAPPA character Agent ("James, the butler") operates in the context of a full screen interface representing an on-line winestore. Product images are presented using a variant of the "Rapid Serial Visual Presentation" method (RSVP) developed by Spence and others (deBruijn and Spence, 2000; Spence, 1999). We utilise the RSVP presentation technique to allow a large range of products to be displayed in a small screen area. It appears to be particularly appropriate when a retailer wishes to present a substantial stock range for immediate viewing, as in a shop environment. RSVP represents an effective, and distinctive, alternative to the more conventional and common-place "catalogue pages" approach to on-line site design. MAPPA also acts as a test platform for a variety of other Agent based technologies. Each component part of the system is encapsulated as a FIPA-compliant software agent, which inter-communicate using the FIPA-ACL Agent Communication Language.

Much work has been done to integrate visual character Agents ("synthetic personae") into electronic display systems, notably in entertainment, CAE and e-retail contexts. Despite the number of systems that have been proposed and developed, the number of empirical evaluations of users' reactions to these characters is still relatively small and their interpretation can be problematic (Arafa *et al*., 2001; Massaro *et al*., 1998; McBreen *et al*., 2000; Van Mulken *et al*., 1998; see Dehn and van Mulken, 2000 for an overview). Arafa *et al*. (2001) have performed analyses based on user perceptions of the MAPPA system and found that users preferred the MAPPA interface when it included the character Agent together with text and voice compared to text-and-voice and voice-only interfaces. The MAPPA character Agent was also judged to possess a number of positive social characteristics. In this study we extend this work by focussing on the extent to which the MAPPA character Agent is capable of attracting and holding the user's attention. Here we use eye-tracking equipment, which allows us to record where a user is looking on the screen at each moment.

Eye-tracking, or gaze tracking, is an established technique for determining what a person is attending to at any particular time. Eye movements are partially volitional, under the "conscious" or active control of the participant, and partially autonomous, selected by an attention mechanism inherent to the brain's visual system. While eye movement data is generally a good indicator of attention, it represents no guarantee as to the actual object of a person's attention. An individual may, of course, have periods during which they are thinking of something quite unrelated to where their eyes are directed. Nevertheless, it is a powerful tool for measuring attention within a participant that is both quantitative and largely non-disruptive, as would be the case if the experimenter were obliged to interrupt the participant to ask them what they were attending to.

2 The Study

We are using eye-tracking equipment to investigate in detail how each user reacts to a synthetic personae in the context of a particular kiosk based e-retail system. We ask our participants to complete a number of kiosk sessions during each of which the ePSA (James, the character Agent) appears on the screen to present information in an explanatory ("help") context, or to present additional information about products displayed on the screen. This study will concentrate on analysis of the participant's gaze control relative to the appearance of the synthetic character Agent, James, in these two primary roles, as system guide and to present product information during a MAPPA session.

We conjecture that, ideally, the effect of James on gaze control during "help" episodes will be different to product presentation episodes. During "help" episodes it is appropriate that the participant's visual attention is drawn to the screen items for which help is being provided. During presentation episodes it is appropriate that the participant's visual attention is focused on the product item being considered.

We isolate these Agent related episodes and analyse the data obtained for the ratio of visual (apparent) attention paid to (a) the ePSA, (b) the screen component being described, and (c) other screen objects not directly related to the ePSA's activity. This data may be analysed along three dimensions:

(1) The variability between individual participants
(2) The variation over time with repeated exposure to the ePSA by individual participants
(3) Any correlation between the captured data and the participant's comments about the ePSA

We believe this will provide an important addition to previous studies, and provide a method to qualify studies that otherwise rely heavily on subjective responses by users. We are also interested in the effect of the RSVP display on visual attention, but this is not the focus of this investigation.

3 Eye-tracking

The movements of the human eye are not smooth, but proceeds as a series of *saccades*, rapid shifts of the gaze from place to place and then holds steady for short periods called *fixations*. Saccades are usually considered to be ballistic, that is the brain chooses a place to look, and the gaze is directed there without any further control. Saccades typically take between 30-120ms and can cover up to 40 degrees of visual angle. It is widely believed that little or no visual processing takes place for the duration of a saccade. At the end of each saccade a brief period of correction appears to take place to align the fovea over the point of interest. Saccades are characterised by periods of high acceleration and rapid angular change, fixations by periods (typically between 200 and 600ms) of low eyeball acceleration.

The gaze may track a moving point of interest (for instance, the track of a mouse pointer) at a speed quite distinct from a saccade, however all movements across a stationary visual field follow the saccade/fixation pattern. Particular tasks, such as reading text, are characterised by sequences of small saccades and short fixations across the extent of the text. When the head is rotated eye movement takes on a sawtooth like motion, referred to as *nystagmus*. Participants are required to keep their head still throughout our experiments, so we would not expect to encounter this phenomenon. However, participants blink frequently during an experiment, and this causes momentary loss of tracking data.

Eye-tracking equipment has been available for many years and has been used for a wide variety of purposes. Devised to formulate and test theories of human perception (for example, Binello, *et al*. (1995), Scialfa and Joffe (1998)), eye-tracking equipment has subsequently found practical application in the evaluation of web page design (for example, The Stanford Poynter Project, 2000), and as a way to evaluate the effectiveness of advertising material (Rosenbergen, 1997). Hodgson *et al*. (2000) have used eye-tracking to evaluate visual planning strategies in the "tower of London" task. There have been several proposals for gaze contingent systems (notably Starker and Bolt, 1990), where the appearance of the screen changes according to the point of gaze, and for using eye movements as an control selection device, complementing existing methods (for example, Jacob, 1995). The equipment we use here is normally employed by a major London teaching hospital for clinical evaluation of patients with visual impairments and (also for) neurological research.

4 The MAPPA Interface Screen

Figure one shows the layout of the MAPPA interface. This interface has been through several design iterations and user evaluation stages, and represents a typical trade-off between providing clear and easy access to functions the user requires against excessive screen clutter. Two items of particular note are the RSVP display area (the area of the screen containing the bottle images, item (1) on the figure) and "James", the ePSA, who appears on the screen to present information to the user (2). In general, the ePSA does not appear unless the user takes some specific action, such as click on a bottle or select a "help" button.

The RSVP display is intended to emulate a "shelf" of products. The user may rapidly move back-and-forth along the shelf using the tracker-slider control (3) or left/right buttons (4) and (5). Having found a part of the "shelf" with products of interest, the user may "riffle" through those products by running the mouse pointer over them. As each product is touched by the pointer it raises slightly and a brief product description immediately appears above it, emulating a shelf product label (7). James, the ePSA, may be called to provide further information by clicking on a product image. The RSVP effect, which is both distinctive and visually effective, is not easily envisaged from the still image of figure 1. The effect has garnered unanimous interest whenever demonstrated informally or presented under controlled test conditions (Arafa, *et al*., 2001).

The ePSA, which is based on Microsoft's Agent technology, is animated over the top of the display. The Agent's utterances are generated in a variety of ways, some scripted within the control program, some derived from an extensive product information database, and can be generated by a JESS based expert system Agent, which has access to both general product (wine) knowledge, and user preferences collected during the current and from previous sessions. Information presented by the ePSA is both vocalised using a male voiced electronic synthesiser, and appears in a scrolling "speech bubble" adjacent to the Agent's head (item (6) and figure 2 – detail).

Initially there may be several hundred products on the display "shelf". The user may use the drop-down boxes, or call up product refinement dialogs. This is typically refined to a small number of products (perhaps 5-15) before final product selection is made. This is described in the next section.

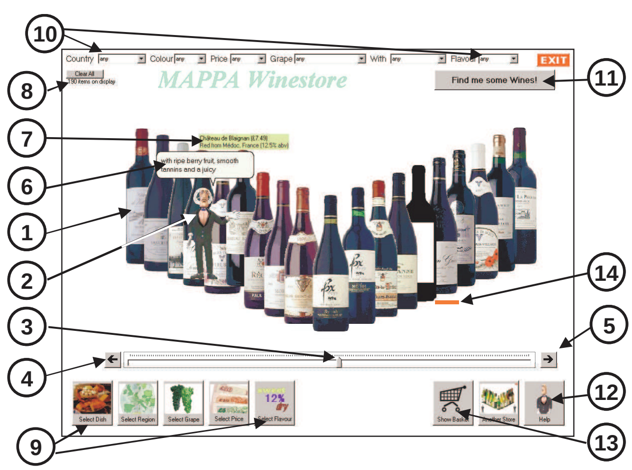

Figure 1: The MAPPA display screen

Figure 2: detail of James, the ePSA

4.1 Refining the Product Selection

This section describes the two ways in the current display area by which the user can directly request sub-sets of the full product range by explicitly applying selection criteria. First the user may make selections from the series of "drop-down" selection boxes (10). Alternatively, the user may initiate detailed selection dialogs from button based selection panels (9). These dialogs appear at the top of the screen, and normal operation of the RSVP screen is suspended while they are displayed.

The first way is intended to reflect the method available from most web-sites, each drop-down allowing a number of broad categories. The second (dialog based) way allows for highly detailed or specialised options to be selected. Figure 3 shows the detailed "goes-with" selector. The user may refine a selection of wines on the basis of food it might best accompany. The main graphic buttons on this dialog access a broad category of dish, reflecting the categories available in the first method. The list box selector within this dialog (whose contents depend on the button selected) allows for a detailed choice to be made. Equally, the user might refine their choice on the basis of region, price, flavour or grape variety, each having its own dialog, each of broadly similar design.

Figure 3: The "goes-with" selector dialog

Figure 4: Expressing Preferences

Note, in particular the "Ask James" button on this selection dialog (an equivalent button is placed on the other selector dialogs). This option invokes the appearance of James, who will explain the purpose and use of each facility associated with the dialog.

The Personal Service Agent could update the product selection at any time, but the "Find" button (11) is intended to allow the user to request the PSA to present a new selection based on the criteria they have specified. Criteria may be selected sequentially, allowing the user to cumulatively refine the product search at successively greater degrees of detail.

4.2 Expressing Preferences

In general, we prefer not to present the customer with an extensive "preferences" form to be completed, but rather seek to gather that information incrementally by fully elective means. Each product selection dialog presents two extra buttons, "Like" and "Dislike", figure 4, shown circled. Pressing either of these buttons informs the Personal Service Agent that the customer prefers (or does not) the current selection that is indicated. This information then forms part of the persistent knowledge the ePSA holds about the customer. The strength of a particular preference (or dislike) may be indicated by repeated application of a button and previously held preferences reversed using the opposite sense button.

Many of the product selector panels allow the user to select a broad category (i.e. "France") with a pictorial button, or (for more knowledgeable users), to specify a detailed sub-category (such as "France-Bordeaux") from the list-box. The preference always attaches to the specific option selected. In this manner the user may express a general dislike of a general category, such as sweet wines, but a liking for exceptions, such as the sweet Tokaji style, or even an individual product (using the preference buttons attached to the appropriate dialog).

4.3 Other Dialogs

During a complete MAPPA session the user will be exposed to five other, specialised, dialogs. First a "login" screen, where the user is required to enter a user name to access the system (and so be identified to the system on subsequent visits). Second, a registration screen, which appears directly after login on the first visit only. It requests essential user details, such as delivery address. The ePSA appears with these two dialogs to explain what is required. Directly after user details screen, the user is given the option to view a brief "help" sequence, during which James describes the major components of the system. This may be invoked at any time using the "Help" button (12). Third, the user may call up a purchase selector (by double clicking on a product image). Fourth, the user may view and hide items already selected for purchase in the "Basket", button (13). Lastly a checkout dialog, which prompts for credit card details to complete the transaction. This last dialog has, of course, no effect.

5 The Eyelink Equipment

These experiments use the "Eyelink" eye-tracking equipment from SR Research (2000). Eye movements are detected using two cameras facing towards the participant's eyes. This arrangement is shown in figure 5. A third camera faces forward toward a pattern of LED emitters attached to the side of the display screen (visible in figure six) to detect movements of the participant's head, and can correct the calibration when this happens. Eyelink measures the eye position and calculates the X and Y screen coordinates for the participant's gaze every 4ms. The equipment also detects and calculates the start and end times, and start and end coordinates of each saccade, fixation and blink event (to the nearest millisecond) and stores these events in a FIFO buffer. Eyelink events are transferred to and recorded on a second computer running the MAPPA system over an Ethernet link.

We recover the fixation events, and the current position of the eye coordinates from the FIFO buffer every 50 ms, and record this data, along with mouse coordinates and every significant event generated (user initiated actions, dialogs and activity by the ePSA) into a trace file for subsequent analysis. Simultaneously we record a digital videotape of the screen display, along with the sound of the ePSA and any comments by the participant or conversation between participant and experimenter. The video signal is derived directly from the display using the “dual-head” facility on a Matrox G450 graphics card. This video sequence will later be rendered down to a “.avi” video clip and correlated with the data from the trace file.

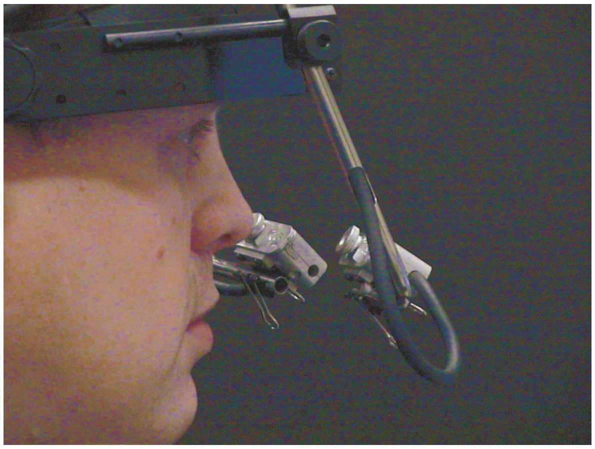

Figure 5: Eyelink sensors

Figure 6: Eyelink in operation

During a typical session with the ePSA character Agent the user’s focus of interest on the screen changes often and rapidly, for instance when the character Agent appears, or when some activity must be performed.

6 Experimental Procedure

Our approach is to involve each participant in a complete MAPPA session while connected to the Eyelink recording apparatus. The session starts with the login screen, and will end with a (simulated) purchase using the checkout screen. Each participant is taken through a set procedure by the experimenter, who acts as “facilitator” during each session. The facilitator may prompt the participants to perform various activities, to ensure the appropriate number of Agent interactions were captured within the overall session structure.

We obtained experimental data from 5 participants who volunteered to take part in this study. None of the participants used in this pilot study had used MAPPA before, but some were familiar with the Eyelink equipment. All participants were given a brief description of the MAPPA system and an indication as to the purpose of the experiments. Participants who had not used Eyelink before were introduced to the equipment and purpose of the experiment while the equipment was attached to their head. All participants then had to complete a calibration sequence (during which the Eyelink establishes its coordinates for the MAPPA display screen). Data capture and video recording start once calibration is complete and the login screen appears. All the participants completed the procedure at least once and three of the participants twice.

On their first exposure to the MAPPA system participants complete the registration procedure (about 60 seconds). For these experiments registration was truncated to entering their first name only. In earlier tests we noted that typing at the keyboard disrupted the Eyelink calibration, so the facilitator typed in the participant’s responses. The participant retained the mouse, and the keyboard is not used after the registration screen.

When prompted whether they would like to view the introductory help sequence the facilitator prompts them to do so. This sequence lasts approximately 41 seconds and involves the ePSA moving to five distinct locations on the screen while describing the functions and facilities of the MAPPA system. Once this is completed the user is allowed a short time to explore the interface. During the session the facilitator will prompt the participant (if they do not do so spontaneously) to do the following things:

1. Use the “left” and “right” buttons and the tracker bar to view significant portions of the product “range”.
2. Invoke the ePSA to provide product information (by clicking on a product image) between 5 and 10 times.
3. Call at least one product selection dialog, and invoke the help sequence attached to it (the ePSA provides help at five distinct locations on the screen, lasting approximately 48 seconds).
4. View at least one reduced product selection formed as a result of using the dialog.
5. Use the purchase dialog to add at least one product to the shopping basket.
6. View, and hide, the shopping basket.
7. Invoke the Checkout dialog and exit normally.

The overall intention (un-stated to the participant) behind this procedure is to capture at least 15 distinct presentations by the ePSA, embedded in a sequence of other activities. At least 10 of these instances are where the ePSA provides help or advice, and between 5 and 10 instances where the ePSA provides product information. The ePSA also makes anumber of other appearances, for instance to welcome the user, or to indicate the results of a product selection search. These are analysed in a separate category. While a great deal of other information is also captured, it forms no part of the analysis reported here.

Data recording (i.e. actually using the Eyelink apparatus) lasts between 6 and 10 minutes, this is both to restrict participant fatigue and keep the data captured to manageable proportions. It also reflects the time taken to complete a "normal" interaction. Finally the facilitator and participant "de-brief", the facilitator noting any comments or observations about the ePSA or MAPPA offered by the participant. Participants are invited to return and complete further sessions. Subsequent session are organised as the first, except that the MAPPA system recognises the user, omits the registration screen, moving immediately to system use.

7 Data Analysis

We have been obliged to develop an analysis tool, allowing the experimenter to analyse the data on an event-by-event basis and to correlate these events directly to individual frames in the video sequence. The analyst may step through the trace file to identify each eye-movement, fixation and mouse movement. The tool also reports the text of the most recent action invoked for the ePSA and each user response or activity.

The principle analysis required for this study involved identifying the instant at which the Agent appears on the screen, and defining a range of events upto and including the instant the Agent disappears from (or moves to another location on) the screen. All Eyelink events in the defined range may be viewed superimposed over a (selected) background frame. Once a range is presented in this form, the tool will calculate the number of fixations present, the overall time, the mean time and standard deviation for the fixations shown. The analyst may then "lasso" around various clusters of fixations, such as those over the speech bubble, over the ePSA, or over the product image. The tool then calculates the same values for the selected fixation events as a proportion of those for the complete displayed range.

Figures 7 and 8 show examples of a user's visual attention during individual ePSA appearances. This information represents the raw data for our analysis. Each fixation is represented by an 'F' (the top, left corner representing actual screen coordinates of the gaze point), the joining lines the saccades. It is easy to determine the sequence and relative "interest" shown in each of the screen components. The analyst may easily swap between cluster view mode and tracking individual events to fully understand the sequence of eye movements made by the participant.

We noted that the Eyelink apparatus is particularly prone to abrupt shifts of calibration whenever the participant moved their head (despite the compensation mechanism described previously). Whilst every participant was enjoined to keep their head still throughout the session, not all appeared to be able to do so. Prior trials with various schemes for head restraint or chin rest proved either ineffective, or unacceptably uncomfortable, and had to be abandoned. The analysis tool provides a re-calibration facility. Fortunately the shifts are distinctive, and the data may be effectively re-based through the session.

Figure 7: Eyelink data overlaid the ePSA

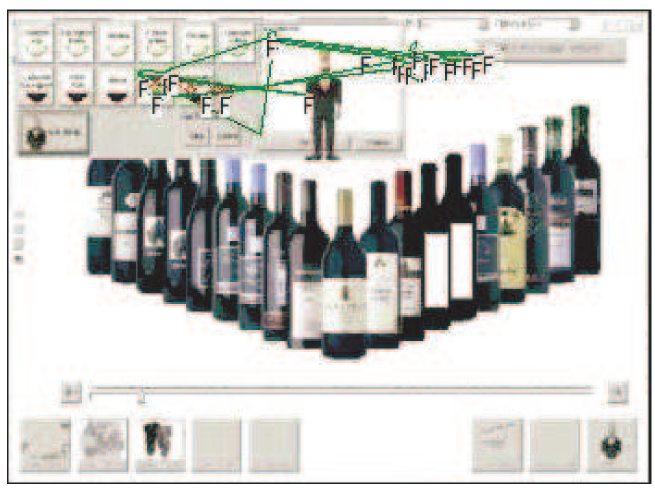

Figure 8: Attention sharing

8 Results

The results reported in this paper are of work in progress. Our aim here is to show that gaze tracking provides a useful addition to the methods used to evaluate the effects of character Agents in the user-interface. In particular we note the extent to which the presence of the character Agent attracts and holds the participants visual attention for the period it is on-screen, the extent to which this varies between individual participants, and the degree to which it changes with subsequent exposure to the system. Figure 9 shows a typical course of events during an ePSA appearance. It shows the participant's gaze (indicated by the Fs) shifting from a selected bottle to the description of the wine in that bottle appearing above, then to the Agent's face when it first appeared in front of the second bottle on the left and finally to the Agent's face when it moved to the location it is seen at now. This snapshot was taken 1.072 seconds after the Agent first appeared on the screen. The first fixation on the Agent occurred 644ms after its onset and again 248ms after its move. To confirm that this example is not just a one-off result we have measured the time between Agent onset and first fixation on the Agent for 77 Agent appearances and found that on 50 of those occasions the first fixation occurred within 1 second.

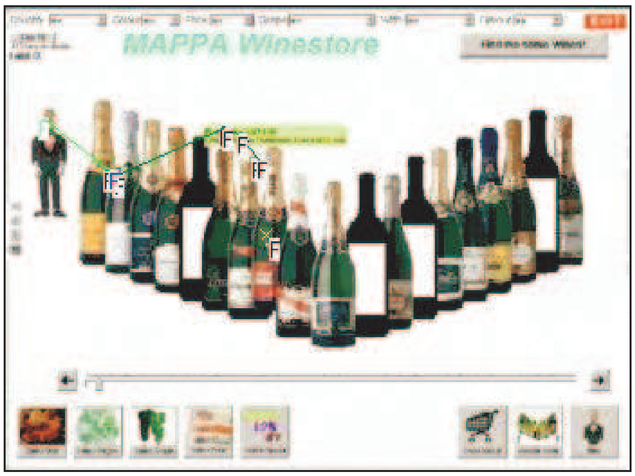

Figure 9: Attention grabbing

In addition to information about the attention capturing properties of the MAPPA ePSA character Agent, the Eyelink data may also provide information about the characters' hold on users' attention. We are particularly interested in contrasting participants' engagement of attention during different types of roles that the ePSA is capable of taking on. We are also interested in comparing the role in which the Agent acts as a guide to the MAPPA interface (Interface Guide) with the one in which the Agent gives the user advice about wines (Wine Adviser).
For the purpose of analysing what participants were attending to during ePSA appearances, we divided the fixations into five mutually exclusive categories depending on the interface object receiving the fixation. This classification was based on the relevance of the underlying interface object to the Agent's role. Fixations during the Wine Advisor and the Interface Guide appearances were divided into these five categories:

1. those directed towards the Agent interface character (Agent) ,
2. the speech bubble (Text),
3. the bottle about which the Agent is advising (Product)
4. other interface objects not related to the Agent's advice (Roaming).
5. Any fixations that could not be classified as belonging to any of the categories above were scored as Miscellaneous.

Table 1 presents a summary of the fixations for the two agent mediated roles in question. The numbers represent percentages of the total fixation time during the Agent appearances. For example, from the table we can tell that subject #1 directed 44.5% of his gaze during the Wine

Table 1: A summary of the data obtained during each participant's first MAPPA session. The data represent percentages of the total gaze time during all of the Wine advisor and Interface Guide episodes. N indicates the number of Agent episodes on which the percentages were based.

Wine Advisor					
	#1 N=6	#2 N=8	#3 N=7	#4 N=6	#5 N=7
Agent	38.9	16.4	28.6	14.5	9.3
Text	44.5	67.1	45.4	74.3	75.5
Product	12.9	12.4	13.5	8.9	11.1
Roaming	3.7	1.8	12.5	0.0	3.8
Misc.	0.0	2.2	0.0	2.3	0.3
Interface Guide					
	#1 N=10	#2 N=10	#3 N=10	#4 N=10	#5 N=10
Agent	21.1	23.0	13.4	12.2	19.9
Text	45.9	43.0	36.1	65.4	61.9
Product	11.2	13.7	5.2	6.6	6.5
Roaming	19.9	13.4	42.6	9.4	8.5
Misc.	1.9	6.8	2.6	6.5	3.3

Advisor episodes towards the speech bubble (Text). The data in Table 1 suggest that most of the fixations during both types of Agent episodes are directed towards the speech bubble. From the pattern that is generally observed for these particular fixations (see for example Figure 8) we can conclude that this indicates that approximately half the duration of the Agent episodes participants are engaged in reading the text appearing in the speech bubble. Moreover, it appears that hardly any time is spent looking at either the product or its description appearing above the bottle. It also appears that during the both types of episodes participants spent almost the same length of time, if not more, looking at the other objects in the interface. Most importantly, however, it seems that there is little difference in the pattern of fixations between the two types of Agent roles for any of the participants. If anything, it appears that participants may spend a little more time roaming the interface during the Interface Guide episodes.

The data presented in Table 1 indicate that participants' attention is substantially engaged by the Agent's visual character. One of the assumptions underlying the use of these synthetic character Agents is that they allow users to bring to bear their skills in social interaction. When people engage in person-to-person interaction a lot of time is spent attending to the other person's face because valuable information can be gleaned from lip movement and facial expressions. Similarly, body posture and gestures can be used to engage and direct user's attention. If these assumptions underlying the use of character agents have any validity one would expect participants to focus their attention predominantly on the Agent's face and hand gestures. Therefore, we further analysed the fixations classified as Agent in Table 1 to see whether participants' fixations on the character agent were clustered on the Agent's face and gestures or distributed across the Agent's body randomly.

Table 2: Percentages of the time looking at the Agent character engaged in looking at its face, gestures or body for each of the participants' first session

	Agent parts		
Ps	Face	Gesture	Body
#1	83.0	13.2	3.8
#2	88.1	2.3	9.7
#3	74.5	13.2	12.3
#4	63.9	16.0	20.1
#5	73.7	20.8	5.5

The data presented in Table 2 shows how much of the total time engaged in looking at the Agent is spent looking at the Agent's face, its gestures and its body. From this data it is immediately clear that when participants focus their attention on the Agent, they are looking mostly at the Agent's face. The Agent's gestures do not seem to attract a similar amount of attention. However, the fact that the fixations on the agent are not distributed randomly across the agent's parts suggests that participants do interact with a character agent in a social manner.

The third question we were seeking to answer was whether participants distribute their attention across the interface during the Agent episodes differently when they use MAPPA the second time compared to the first time. In order to answer this question we asked three of the five volunteers to come back for a second MAPPA session approximately 10 days after they took part in the first session. The data for these participants' second session is shown in Table 3. Only data for the Wine Advisor episodes are shown because, as the participants were already familiar with the MAPPA interface, the assistance of the Interface Guide was not requested during session two. Comparing the data in Table 3 with those of the Wine Advisor in Table 1 reveals that the participants may have spent slightly more time reading the text in the speech bubbles, possibly, at the cost of spending less time looking at the agent. Participant 1 appears to have spent a bit more time roaming around the interface.

Table 3: Percentages of fixation time directed to the various interface objects during Agent episodes during the second session.

	#1	N=5	#2	N=7	#3	N=7
Agent	16.9		14.0		18.5	
Text	54.0		70.6		61.9	
Product	7.1		14.0		3.5	
Roaming	13.1		0.8		15.1	
Misc.	8.8		0.6		0.9	

9 Discussion and Conclusions

Several studies have addressed the question whether a character Agent attracts and holds participants' attention. When asked, participants have report that character Agents do attract attention (Koda & Maes, 1996; van Mulken *et al.*, 1998) and that they do not distract from the task at hand more than non-character representations of Agents (van Mulken *et al.*, 1998). This attention capturing quality of character Agents has been confirmed in the present study using eye movement data. However, the focus on the character Agent may have distracted from the other objects in the interface, most notably the products and objects about which the agent is providing information. Our findings are also in line with those of an earlier study by Takeuchi and Naito (1995) into the attention capturing quality of character agents using gaze tracking, which showed that a character agent captures attention to a larger extent than the same agent replaced by an arrow.

In summary, several lines of evidence suggest that character agents possess an attention capturing quality not possessed by other non-character representations. It remains unclear, however, whether this quality of character agents is beneficial towards users' ability to interact with such systems easily and naturally in order to achieve their goals. This is one of the issues we intend to investigate in future research using gaze tracking data.

Gaze tracking represents a powerful and effective tool in the analysis of interface usage. The technique provides a measurable and direct insight into what an experimental participant is attending to when a character Agent appears on the screen. The experimental regime established allows for detailed data capture and subsequent analysis of this information. The regime allows direct and quantifiable measurement to be made of the variability between participants and the changes in a participant's responses with repeated exposure to the Agent. It also provides a benchmark to evaluate user's (necessarily) subjective comments about character Agent interactions.

Acknowledgements

This work has been partially funded by EPSRC Grant Number GR/N20881. The MAPPA system was originally developed under a grant from the European Commission ESPRIT projects funding, under grant number EP28831. Special thanks are due to the Department of Neuroscience and Psychological Medicine at Charing Cross Hospital for access to the Eyelink equipment and in particular to Dr. Tim Hodgson for his help and expertise.

References

Arafa, Y., Dionisi, G., Mamdani, A., Pitt, J., Martin, S. and Witkowski, M. (2000) "Towards Building Loyalty in e-Commerce Applications: Addressing Issues on Personalisation, Persistence & Presentation", 4th Int. Conf. on Autonomous Agents, Agents in Industry Workshop, 8pp.

Arafa, Y., Fehin, P., Pitt, J. and Witkowski, M. (2001) "Indicators to the Effect of Agent Technology on Consumer Loyalty", to appear.

Binello, A., Mannam, S. and Ruddock, K.H. (1995) "The Characteristics of Eye Movements Made During Visual Search with Multi-element Stimuli", Spatial Vision, Vol. 9, No. 3, pp. 343-362

deBruijn, O. and Spence, R. (2000) "Rapid Serial Visual Presentation: A space-time trade-off in information presentation", in proc. Working Conference on Advanced Visual Interfaces (AVI 2000), pp. 189-192

Dehn, D.M. and van Mulken, S. (2000) "The Impact of Animated Interface Agents: a Review of Empirical Research", International Journal of Human-Computer Studies, 52, 1-22

Hodgson, T.L., Bajwa, A., Owen, A.M. and Kennard, C. (2000) "The Strategic Control of Gaze Direction in the Tower of London Task", Journal of Cognitive Neuroscience, Vol. 12, No. 5, pp. 894-907

Jacob, R.J.K. (1995) "Eye-tracking in Advanced Interface Design", in: Barfield, W. and Furness, T.A. (eds.) "Virtual Environments and Advanced Interface Design", New York: Oxford University Press, pp. 258-288

Massaro, D.W., Cohen, M.M., Beskow, J., Daniel, S. and Cole, R.A. (1998) "Developing and Evaluating Conversational Agents", in: Proc. Workshop on Embodied Conversational Agents (WECC-98), Lake Tahoe, 11pp

McBreen, H., Shade, P., Jack, M. and Wyard, P. (2000) "Experimental Assessment of the Effectiveness of Synthetic Personae for Multi-modal E-Retail Applications", in: Proc. 4th Int. Conf. on Autonomous Agents (Agents 2000), pp. 39-45

Pitt, J. Arafa, Y., Dionisi, G., Martin, S. and Witkowski, M. (2000) "Creating Loyalty in E-Commerce Using Agent Technology" e-2000

Rosenbergen, E., Wedel, M. and Pieters, R. (1997) "Analysing Visual Attention to Repeated Print Advertising using Scanpath Theory", SOM Research School report #97B32, University Library Groningen.

Scialfa, C.T. and Joffe, K.M. (1998) "Response Times and Eye Movements in Feature and Conjunction Search as a Function of Target Eccentricity", Perception & Psychophysics, Vol. 60, No. 6, pp. 1067-1082

Spence, R. (1999) "A Framework for Navigation" Int. J. of Human Computer Studies, Vol. 51, No. 5, pp. 919-945.

SR Research/Eyelink (2000): http://www.smi.de/el/elsys.htm

Starker, I. and Bolt, R.A. (1990) "A Gaze-responsive Self-disclosing Display", in: proc. ACM Human Factors in Computing Systems Conference (CHI-90), pp. 3-9

Takeuchi, A. and Naito, T. (1995), "Situated Facial Displays: Towards Social Interaction", in: I. Katz, R Mack, L. Marks, M.B. Rosson & J. Nielsen, Eds. Human Factors in Computing Systems: CHI'95 Conference Proceedings, New York: ACM Press, pp. 450-455

The Stanford Poynter Project (2000), http://www.poynter.org/eyetrack2000/

Van Mulken, S., André, E., and Müller, J. (1998) "The Persona Effect: How Substantial Is It?" in: proc. HCI-98, pp. 53-66