



University of
Salford
MANCHESTER

Evaluation of anthropomorphic feedback for an online auction and affordances

Murano, P and Holt, PO

Title	Evaluation of anthropomorphic feedback for an online auction and affordances
Authors	Murano, P and Holt, PO
Type	Article
URL	This version is available at: http://usir.salford.ac.uk/2727/
Published Date	2011

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.

Evaluation of Anthropomorphic Feedback for an Online Auction and Affordances

Dr Pietro Murano¹ and Prof Patrik O'Brian Holt²

¹ School of Computing, Science and Engineering, University of Salford, Salford, M5 4WT, UK

² School of Computing, The Robert Gordon University, Aberdeen, AB25 1HG, UK

Abstract

This paper describes an experiment investigating the effectiveness and user satisfaction of using anthropomorphic feedback at the user interface. The context chosen was online bidding due to this kind of activity being very much used in current times by general users. The main results of the experiment were that there was a statistically significant effect observed for the time taken to place a bid in the anthropomorphic text condition. However there were no other significant effects for effectiveness issues and user satisfaction indicators. The results were also analysed in terms of the affordances and the main findings were that each of the four conditions tested in the experiment were probably equivalent in terms of their facilitating the affordances. Overall it may be more important to facilitate the affordances rather than a type of feedback being anthropomorphic in nature or not.

Keywords: *Anthropomorphic, affordances, user interface evaluation.*

1. Introduction

For several years research into anthropomorphic user interface feedback has been conducted. However, despite many claims in the literature, there is no overall pattern of results that could aid user interface developers to make a good informed choice for the use of such feedback in a software system.

Therefore, the main objective of this research is to aid in the improvement of user interfaces by better understanding the effects of using anthropomorphic user interface feedback. Specific concentration is placed on comparing anthropomorphic and non-anthropomorphic user interfaces to address the issues of effectiveness and user satisfaction in relation to context and domain and to provide some explanation of the results in terms of an appropriate theory such as the theory of affordances.

There are various opinions amongst the computer science community regarding the effectiveness and user

satisfaction of anthropomorphic feedback at the user interface. Some researchers are in favour of anthropomorphism, e.g. Koda and Maes [10], Maes [13], Laurel [11], Agarwal [1], Zue [33] and Takeuchi and Naito [30]. However, some researchers are not generally in favour of anthropomorphism in most circumstances e.g. Shneiderman and Plaisant [29]. Each of these researchers tends to base their opinions on various studies conducted in the area. As stated, due to the inconclusive nature of the results of these studies, there is the need for more work in this area to gain a better understanding of such differences in opinion and experimental results.

The rest of this paper will consider some key previous research, then the experiment conducted will be described in detail along with the experimental results and conclusions. Further, the results will be discussed in the context of affordances.

2. Some Key Previous Research

This section will aim to discuss research which has already been carried out by others and the authors of this paper on anthropomorphism and highlight some of the differences in results obtained by other researchers. Research about anthropomorphism spans various contexts including agent-based software.

The first paper to consider was an experimental study by Moreno et al [16] in the context of tutoring and learning about plants, they found that experimental participants using an anthropomorphic agent were better able to use their newly learned knowledge to solve similar problems in the same domain. They also found that participants in this group had more motivation to continue learning about plants and had overall more interest in the subject area. No difference was found for actual memory capacity. The control group used the same information as in the anthropomorphic agent group, but the agent was substituted with text.

Another study in the area of tutoring by Moundridou and Virvou [15] tested 2 conditions in an algebra tutoring environment. The first condition had a talking synthetic face and the second was the same as the first condition with text replacing the synthetic face. The main results showed that there was no significant difference between the 2 conditions for task time completion. However the participants in the anthropomorphic condition enjoyed the experience more, found the system more useful and less difficult to use.

The studies by Moreno et al [16] and Moundridou and Virvou [15] were in similar contexts, but the results were not the same. Moreno et al's [16] study showed more positive outcomes in favour of anthropomorphic feedback compared to the study by Moundridou and Virvou [15]. Although it is acknowledged that the measures used in each study were slightly different and could be a factor. However the results for user perceptions were similar, where the anthropomorphic feedback fostered more positive user perceptions.

In another study by Catrambone et al [3] an experiment was conducted using an editing environment and a travel items recommendation environment. Three conditions were tested with these two environments. The first condition was an animated agent, the second was a still photograph of the same agent used in the first condition and the third was a cartoon image of a lit light bulb. The tasks involved doing some editing in an unknown word processor and making some choices regarding what items to take on an international trip. During the editing task the agent was reactive in nature, while in the travel items task the agent was proactive. The main results were that for the travel task, the participants were generally influenced by the agent's suggestions. However no effect was recorded for type of agent. In the editing task, there was no difference in task time across the 3 conditions. Further, the participants felt that the agent was less intrusive and more worthwhile in the editing task than in the travel task. Also the participants were observed to be at ease in querying the agent for help in the editing task while the converse was true for the travel items task.

Furthermore in a related study by Xiao et al [32] an experiment was conducted in an editing environment testing three experimental conditions. The first was a reactive anthropomorphic agent, the second was also an agent that was reactive and proactive in nature and the third was a control condition consisting of an approximately equivalent paper based manual. The main results for the experiment show that there were no significant differences in task time and number of

commands used across the 3 conditions. After the experiment, participants were also asked to recall as many editor commands as possible. This aspect did not produce any significance across the 3 conditions. There were also no significant results in the participant opinions about the agents and paper manual.

The studies by Catrambone et al [3] and Xiao et al [32] were similar in nature, but the results clearly differ, where Catrambone et al's [3] study shows some results favouring anthropomorphic feedback, while Xiao et al's [32] study shows no differences overall in the experimental conditions used.

In a study by David et al [5], the authors conducted a three condition experiment in the context of a quiz about ancient history. They were investigating different anthropomorphic cues in terms of character gender and attitude and user perceptions about the character in relation to quiz success (or not). The overall results of their experiment suggested that anthropomorphic cues led to users believing the character to be less friendly, intelligent and fair. This finding was linked with the male character and not with the female character.

However, in a study by Qiu and Benbasat [28] a six condition experiment was conducted to test 'perceived social relationship' whilst interacting with a system, in the context of digital camera recommendations. The experimental conditions used in the experiment were animated face with text output, animated face with text-to-speech (TTS) output, animated face with human voice output, no face with text output, no face with TTS output and no face with human voice output. The tasks were of the kind where users dealt with digital cameras and the system provided recommendations to the user, using one of the conditions listed above. The overall results were that the conditions with an animated face elicited greater perceptions of social presence. A similar trend was observed for conditions using a human voice.

The study by Qiu and Benbasat [28] has some contrasts in results compared to the study of David et al [5]. Qiu and Benbasat report positive findings in relation to animated anthropomorphic feedback with a human voice in terms of 'social presence'. However this is in contrast with David et al's [5] study where anthropomorphic cues were perceived negatively in terms of friendliness, intelligence and fairness. These characteristics could also be viewed as being part of 'social presence' [2]. This therefore indicates a further dichotomy of results, despite the experimental contexts being different.

The discussion so far clearly suggests that using

anthropomorphic feedback in a given context does not guarantee better usability in an application. Clearly the above studies have shown that the results overall in various experiments spanning several years at times show anthropomorphism to be better or worse and in some cases not being any different to conventional type feedback. This pattern of inconsistent results in relation to using anthropomorphic feedback has also been observed in the authors' previous work (see Murano, [17, 18, 19, 20, 21, 22]) on anthropomorphic feedback.

In Murano [20] it was shown that in the domain of software for in-depth learning, anthropomorphic feedback was significantly more effective. The results for user satisfaction were not so clear in terms of significance. This experiment was specifically in the context of English as a Foreign Language pronunciation. However in Murano et al [23], also in the domain of in-depth learning and the context of PC building there was no difference between anthropomorphic feedback and non-anthropomorphic feedback in terms of effectiveness for the tasks. The user satisfaction tended more towards the anthropomorphic feedback.

These two experiments were clearly in the same domain and both had the context of instructing users about a particular subject matter. However the results are inconclusive in terms of researchers or practitioners overall being able to conclude that anthropomorphic feedback is better than non-anthropomorphic feedback (or vice versa) .

Specifically related to this paper, are the results by Murano [18]. The paper investigated anthropomorphic feedback in the context of online factual delivery, using the area of direction finding (way-finding) as the specific context. This paper showed with statistically significant results, that non-anthropomorphic feedback was more effective. The results for user satisfaction were not so clear, but participant preferences tended towards the non-anthropomorphic feedback.

In a further experiment by Murano and Holt [24] in the same domain of online factual delivery with the context of online bidding, the results showed more effectiveness for an impersonal human voice. This was a four condition experiment that tested four different voice types (TTS personal voice, TTS impersonal voice, personal human voice and impersonal human voice). The effectiveness concerned users placing higher bids in the impersonal human voice condition. This was effectiveness in terms of an online bidding business owner preferring overall higher bids being placed by users if they were charging commission based on bid amounts. In this experiment user

satisfaction was inconclusive with no significant differences between the four conditions.

These two experiments are also in the same domain, but with different contexts. However the results are clearly different also suggesting a lack of an overall pattern of results.

Therefore, as mentioned in the introduction, this research is aiming to find more information regarding the usage of anthropomorphic feedback, particularly aiming to discover if such feedback is appropriate in terms of effectiveness and user satisfaction. The research is being done in various software contexts and domains. Overall some of the literature argues that differences in results could be due to high sensitivity in the tasks, measures and contexts used in an experiment. However the authors would argue that while this is a possibility, given the body of research that exists in this area, there should be a clearer pattern emerging regarding anthropomorphic feedback and its usability. Therefore with a lack of a clear pattern the authors are suggesting that there could be other explanations to the central issue of anthropomorphic feedback, its usability, and explaining the plethora of results being observed. One such explanation could concern the facilitating or lack of facilitation of the affordances at the user interface whilst presenting a type of feedback to a user.

This paper therefore investigates the domain of online factual delivery further, describing an experiment set in this domain, using again the context of online bidding to test anthropomorphic feedback. This context was chosen because it is a fairly common activity for users of all kinds to carry out over the Internet and was therefore considered to be useful and realistic, whilst maintaining the theme of the previous experiments conducted ([18] and [24]). As with the previous experiments, effectiveness and user satisfaction were the aspects being investigated.

2. Online Bidding Experiment

2.1 Users

- 120 participants were recruited for this experiment.
- Although gender was not the main issue of this research, the participants were students of computer science. There were 105 male participants and 15 female participants.
- Participants were in the 18-35 age range.

All participants had experience or awareness of online bidding as determined through a recruitment

questionnaire. The participants were all recruited from the university population. Specific details about the participants were then elicited by means of a specially designed pre-experiment/recruitment questionnaire which principally asked specific questions about bidding experience and other online activities, e.g.

Approximately how many times have you used online auctions.

1 – 2, 3-5, 6-10, 11-15, 16-20, 21-30, 31+

2.2 Experimental Design

A between users design was used for this experiment. Participants were randomly assigned to one of four conditions. Randomness was achieved by alternately assigning each participant to one of the four conditions, until all participants had been assigned to a condition. The four conditions were anthropomorphic text, anthropomorphic text and a synthetic character, synthetic character only and a control group consisting of only neutral text. The synthetic character had a TTS voice which as in the previous experiment was a generic electronic voice.

2.3 Variables

The independent variables were (1) the types of feedback (Anthropomorphic Text, Anthropomorphic Text and Synthetic Character, Synthetic Character and Neutral Text) and (2) Type of Task (i.e. bidding on three different household items), where the values from the bids made were averaged and included in the analyses (i.e. not the tasks themselves).

The dependent variables were the participants' performance in carrying out the tasks and their subjective opinions.

The dependent measures were that the performance was measured by examining the average bid amount. From an auction point of view, the higher the bid made, the better the business outcome. From a user's perspective the lowest amount is the best outcome. This approach was the same as the one described in Nass and Brave [25], Huang et al [9] and is suitable because it is directly related to the bidding process, which is the main aspect of concern. Further it was decided to put the participants under pressure in the interaction by asking them to bid as quickly as possible to maximise their chance of obtaining the items. The bids and the time taken to place a bid were all recorded automatically by the prototype software. The subjective opinions were measured by means of a post-experiment questionnaire conceptually based on the work

of Nass and Brave [25]. The questionnaire had three main sections where responses were made using Likert [12] type scales. These were sections concerning the general user interface, the participants' impressions regarding the way the items were described and the participants' feelings during the interaction, e.g.

Unclear Description of Items							Clear Description of Items	
1	2	3	4	5	6	7	8	9

2.4 Apparatus and Materials

- A laptop running Windows XP with 256 Mb RAM.
- The laptop's own TFT display was used – 14”.
- CSLU Toolkit (2010).
- Borland C++ Builder.

2.5 Procedure and Tasks

The recruitment questionnaire was handed out and the completed forms were scrutinised for participant suitability. The main aspect that was required for this experiment, was to have participants with an awareness of online auctions. Also it was required by implication to have participants with computing knowledge so that the results would not be biased with issues concerning lack of knowledge in using computer systems, which could indirectly affect bidding behaviour. This was assumed though as all the participants were computer science students. Further at the beginning of the recruitment procedure, participants were informed that they would receive a fixed amount of course credits as a reward for their time.

The suitable participants were then contacted by email, where an appointment during the day was arranged.

Upon arrival the participants were welcomed and comfortably seated at the computer. Then they were informed that all the data collected as part of the experiment would be kept confidential, they could leave at any time they wanted and should they not want their data to be used after the experiment had taken place, this was their prerogative.

Then the participants were asked to read the content of a one page web site which briefly presented a scenario of their graduating from university and moving into a home of their own with the need to buy some household items from an auction. Then some information regarding the bidding process was given and lastly the participants were

informed that they would be asked to complete a post-experiment questionnaire.

When the participants indicated that they had read and understood the material, further verbal details were given regarding the auction process along with reiteration of the auction points found on the experiment web site. This was done to ensure the participants fully understood that the auction they would be using was different to the usual 'Ebay' type auction. Participants were also given a briefing regarding the interaction they would have with the system. The following points were therefore verbally communicated:

- There would be three items and only one bid placed for each item.
- Bid as quickly as possible as they were competing with other 'buyers'.
- The system would inform them of how many other bids had been received for each item, along with the lowest amount bid and highest amount bid (this remained the same for all participants).
- Bid an amount comparative to the actual perceived value of the item. That this auction was similar to a Vickrey type auction [31], which encouraged truth telling as the main approach to bidding. (Note: The strategy of using a Vickrey type auction was used to try and balance the bidding behaviours and bidding strategies of expert bidders in 'Ebay' type auctions with individuals who were perhaps less able or less strategic. It had been a concern that there may have been bias in the results if a straight 'Ebay' type auction had been used, as certain 'Ebay' users have developed various bidding strategies).
- The user interface would have a section with the description of the item and another section with a form allowing bids to be typed into respective fields. Once a bid had been typed, this could be then submitted to the system via a 'submit' button.

After the verbal instructions were completed, the actual auction was started by launching the software. The actual descriptions appeared in a section of the user interface – in whichever feedback mode was being used. Using one of the feedback modes being tested, an initial introduction detailing that there were three items for sale was issued. Then the participant was informed that they would receive a description for each item and that they would need to bid at the end of each item's description.

The next stage involved the system issuing the description for the first item. The description gave a guide price and a brief physical description of the item along with the number of bids already on the item and the lowest/highest amounts already bid. Once the end of the description was

reached the participant typed the bid into the appropriate field of the form on the user interface and then clicked the submit button. Once the submit button was clicked, the second description was issued in the same manner as described above. This was done for three items in total. The actual items were a futon, refrigerator and microwave oven.

During the auction process described in the previous paragraph, the software automatically logged the amount of the bid submitted and automatically recorded the time taken to place and submit a bid on each item. The timing did not include reading time, but was started when the participant began their interaction with the software and stopped when they clicked the submit button for an item.

When the three bids had been placed, the participants were asked to complete a post-experiment questionnaire regarding their subjective opinions and feelings about their interaction experience. When this was completed, they were debriefed regarding the experiment. Specifically they were informed that they were not really bidding against others. This had been done to make the event more realistic. Then they were thanked for their time and asked to not tell anyone what they had done. Finally they were reminded that they would receive the promised course credits for their participation. This procedure was followed in the same manner for all the participants.

2.6 Results

The data collected consisted of performance and subjective opinions data. The performance data was the amount bid and the time taken to bid by participants. The subjective satisfaction data collected from the post-experiment questionnaire which used a Likert type scale involved questions in three main categories. These were the general user interface, the participants' impressions regarding the way the items were described in the auction and the participants' feelings during the interaction.

For the general user interface, the actual questions were about the ease of use of the user interface (UI), the usefulness of the UI, the level of satisfaction of the UI, the clarity of the text used in the UI, the comfort of the colours used in the UI and the intuitiveness of the buttons used.

For the category covering the impression of the way the items were described in the auction, the actual questions were about the clarity of the items' description, if it was like a person, if it was friendly, if it was trustworthy, if it was helpful and if it was likeable.

For the category covering the participants' feelings during the interaction, the actual questions were about whether the participant felt relaxed, comfortable, happy, excited and motivated.

These data were analysed using a multifactorial analysis of variance and when significance was found, the particular issue was then subjected to post-hoc testing for confirmation purposes and to isolate differences from the multifactorial analysis of variance, using a Tukey HSD test.

Interestingly, there was only one significant result observed and this is presented below. All the other non-significant results are not presented in this paper for the sake of brevity.

Firstly the mean and standard deviation is presented for the results in Table 1 below:

Table 1: Mean and SD Average Time

<i>Mean</i>	3.1444167
<i>Std Dev</i>	1.3644335
<i>Std Err Mean</i>	0.1245552
<i>upper 95% Mean</i>	3.3910483
<i>lower 95% Mean</i>	2.897785
<i>N</i>	120

For the variables 'Average Time' and 'group' there was a significant ($p < 0.05$) difference. The anthropomorphic text group on average took significantly longer to place a bid on the items being 'sold', compared to the participants in the other 3 conditions. The F-ratio is 3.82* and this significance was confirmed by the post-hoc test (not presented here for brevity). The F-ratio table is shown in Table 2 below:

Table 2: Multifactorial Analysis of Variance

<i>Source</i>	<i>DF</i>	<i>Sum of Squares</i>	<i>Mean Square</i>	<i>F Ratio</i>
<i>Model</i>	6	37.33949	6.22325	3.8177
<i>Error</i>	113	184.20027	1.63009	Prob > F
<i>C. Total</i>	119	221.53976		0.0017

Lastly the data were also analysed by including the participants as a source of random variation to see if these could have caused some effect. The analysis suggests that there were no actual individual participant effects in the data.

2.7 Results Discussion

No significance was found on the actual amounts bid and the experimental conditions being tested. Furthermore the various post-experiment questions designed to elicit 'satisfaction' (or not) responses, did not reveal any

significant differences between the four conditions being tested.

Furthermore it is unexplained why the anthropomorphic text group seemed to take significantly longer to place a bid than the other three groups under investigation. The timing process in this experiment did not include reading time, i.e. the timing was started when the participants began their interaction with the software. If it had included the reading time, this could perhaps have helped explain the observations, however as this was not the case, the matter remains unexplained. A further aspect that could be suggested is that perhaps the timing mechanism during the observation was flawed for this particular group. However this is not the case either, as the timings were done automatically by the software and had been well tested beforehand and at regular intervals between experimental sessions, for consistency and accuracy.

Another issue that could be raised is that perhaps the participants in the anthropomorphic text group were not 'equivalent' to the other participants in the remaining three conditions. This can be suggested by the fact that the standard deviations accompanying the mean times for each experimental group are approximately equivalent for all the experimental groups, except for the anthropomorphic text group, which had a considerably higher standard deviation (actual SD figures not included for brevity). This seems to have occurred despite the fact that measures were taken to ensure that this did not happen. This included using a large enough sample (30 in each group) and screening the participants for certain desirable characteristics. It was desired to have at least participants that had knowledge of bidding etc, and an approximately equivalent level of education (computer science university students) - this was achieved.

3. The Theory of Affordances and the Results

As mentioned above, the authors are suggesting that there may be other explanations for the lack of a clear pattern of results concerning the usability of anthropomorphic feedback. One such explanation could concern the facilitating or lack of facilitation of the affordances at the user interface whilst presenting a type of feedback to a user.

Gibson [7] was the first researcher to systematically study and propose physical affordances. As the affordances in relation to a computer user interface are different to the affordances discussed by Gibson, a detailed consideration of Gibson's theory is beyond the scope of this paper. However, to set the scene, it would be useful to briefly mention what Gibson meant by using the word

'affordance'. Firstly his theory of affordances concerned organisms and their interaction and reaction with some environment. Further, the word 'affordance' is a fabrication on Gibson's part [7]. Secondly, some of his thoughts can be described by means of one of his many examples. He describes a surface which is flat and solid in some way as 'affording' some sort of support, e.g. for an organism to walk. Conversely a surface that is flat but not rigid, such as a body of water, does not 'afford' the act of walking on this surface (for a human). Thirdly, the issues concerning an organism interacting and reacting to some environment are linked with how one sees and perceives aspects of the environment [7].

However, the initial theory of affordances has been reinterpreted for application to user interfaces. Norman [26, 27] and Hartson [8] are the main sources of the reinterpretations, with more lightweight contributions from Gaver [6] and McGrenere and Ho [14], where they started to apply affordances to computer systems and to decompose affordances into different components.

In the authors' opinion the most substantial effort has been conducted by Hartson [8]. He identifies cognitive, physical, functional and sensory affordances. His rationale is that when doing some computer related task, the users are using cognitive, physical and sensory actions. Cognitive affordances involve 'a design feature that helps, supports, facilitates, or enables thinking and/or knowing about something' [8]. One example of this aspect concerns giving feedback to a user that is clear and precise. If one labels a button, the label should convey to the user what will happen if the button is clicked. Physical affordances are 'a design feature that helps, aids, supports, facilitates, or enables physically doing something' [8]. According to Hartson a button that can be clicked by a user is a physical object acted on by a human and its size should be large enough to elicit easy clicking. This would therefore be a physical affordance characteristic. Functional affordances concern having some purpose in relation to a physical affordance. One example is that clicking on a button should have some purpose with a goal in mind. The converse is that indiscriminately clicking somewhere on the screen is not purposeful and has no goal in mind. This idea is also mentioned in McGrenere and Ho [14]. Lastly, sensory affordances concern 'a design feature that helps, aids, supports, facilitates or enables the user in sensing (e.g. seeing, feeling, hearing) something' [8]. Sensory affordances are linked to the earlier cognitive and physical affordances as they complement one another. This means that the users need to be able to 'sense' the cognitive and physical affordances so that these affordances can help the user.

To recap this experiment detailed in this paper had four different kinds of feedback in the bidding context. The results for effectiveness and user satisfaction were inconclusive. Due to the nature of the conditions involved and the fact that the user interface was identical under each of the four conditions, the authors conclude that the affordances should have been the same or similar irrespective of the four different conditions. The four conditions being tested were anthropomorphic text, anthropomorphic text and a character, character only, and a control of neutral text. On the surface it could appear initially that these different conditions would have fostered different affordances, because some of the conditions were textual only and some had a synthetic character. However this may not have been the case. Firstly all four conditions would have satisfied the cognitive affordance aspect. This is because they would all have facilitated 'knowing' or obtaining the necessary information required to be able to place a bid on the items. While the conditions did have different modes of information presentation (e.g. text or character with a voice), given the context and the amount of information being presented, these modes of presentation would have all satisfied the sensory affordances aspect, which deals with some feature of the software facilitating the user 'seeing' or 'hearing' something. It would have perhaps been a different matter if the information being presented was much more complex and lengthier in nature. The physical affordances aspect would have been identical across the four conditions, because the interface being used across the four conditions consisted of elements for presenting the information to the user and identical elements for entering and submitting bids across the four conditions. This argument also applies to the functional affordances supporting the physical affordances. The user interface elements that facilitated the entering and submitting of bids were clear regarding their purpose, particularly as all participants were briefed in the same manner regarding how bids should have been placed using the system. Looking back to the statistical results for this experiment, there were no clear results to give further indications regarding the reasons for the observed phenomena. One significant result was observed which involved the average time taken to place a bid and the anthropomorphic text group. This group took significantly longer to place a bid compared with the other three groups or conditions. It could be argued that perhaps this condition did not foster appropriate affordances compared with the other three conditions. However this argument is flawed, because if this was true, one would expect that the textual control group and/or the anthropomorphic text and character group to be more closely aligned to this result as they were more similar to each other. However this was not the case, their times were actually closer to the

character only group. However as discussed above this result is unlikely to be an issue of the affordances being adversely affected.

4. Recommendations and Future Work

The authors do suggest further work in this area, particularly where it would be a logical consequence of the authors' experiments to develop further experiments where certain experimental conditions were developed to purposely not facilitate the affordances and compare these with other experimental conditions that specifically facilitated the affordances. If the above arguments are correct, then one would expect that effectiveness and user satisfaction would be greater in the experimental conditions facilitating the affordances and the expectation would be that this kind of result would be statistically significant. However, until further results are available concerning the effects of facilitating the affordances at the user interface, it is suggested that the current available evidence concerning the facilitation of the affordances could be heeded in current interface design to aid usability.

5. Conclusions

This paper has presented and discussed an experiment and the main results of the experiment. The experiment was in the context of online bidding and anthropomorphic user interface feedback. The results have been examined in light of the theory of affordances as interpreted by Hartson [8]. The authors argue that each experimental condition did not differ significantly enough with respect to the facilitation of the affordances at the user interface and therefore this is the main reason for not having obtained more significant results indicating differences in the user interface. However, despite this, the work presented in this paper does add new knowledge to the existing body of knowledge in this area. Also the authors suggest that the facilitation of the affordances is probably more important for usability of an interface than the actual presence of anthropomorphic feedback or not.

References

[1] A. Agarwal, "Raw Computation". Scientific American, 281, 1999, 44-47.

[2] F. Biocca, and C. Harms, "Defining and Measuring Social Presence: Contribution to the Networked Minds Theory and Measure", In F.R. Gouveia, & F. Biocca (Eds). Proceedings of the 5th International Workshop on Presence, 2002.

[3] R. Catrambone, J. Stasko, and J. Xiao, "Anthropomorphic Agents as a User Interface Paradigm: Experimental Findings and a Framework for Research", Proceedings of the 24th Annual Conference of the Cognitive Science Society, 2002, p. 166-171.

[4] CSLU Toolkit, Center for Spoken Language Understanding, Oregon Health and Science University, 2010, <http://www.cslu.ogi.edu/ts/>.

[5] P. David, T. Lu, S. Kline, and L. Cai, "Social Effects of an Anthropomorphic Help Agent: Humans Versus Computers", CyberPsychology and Behaviour, 2007, 10, 3. Mary Ann Liebert Inc.

[6] W. W. Gaver, "Technology affordances", Proceedings of the ACM, CHI 91, Human Factors in Computing Systems Conference, April 27 – May 2, 1991, New Orleans, Louisiana, USA, p79-84.

[7] J. J. Gibson, The Ecological Approach to Visual Perception, 1979, Houghton Mifflin Co.

[8] H. R. Hartson, "Cognitive, Physical, Sensory and Functional Affordances in Interaction Design", Behaviour and Information Technology, Sept-Oct, 2003, 22 (5), p.315-338.

[9] A. Huang, F. Lee, C. Nass, Y. Paik, and L. Swartz, "Can Voice User Interfaces Say 'I'? An Experiment With Recorded Speech and TTS", Unpublished Manuscript, 2001, Stanford, California, USA.

[10] T. Koda, P. Maes, "Agents With Faces: the Effect of Personification", Proc. of the 5th IEEE International Workshop on Robot and Human Communication, 1996, 189-194.

[11] B. Laurel, Interface Agents: Metaphors With Character. In: J.M. Bradshaw, ed. Software Agents, 1997, MIT Press, London, 67-77.

[12] R. Likert, A Technique for the Measurement of Attitudes, 1932, Columbia University Press, NY.

[13] P. Maes, "Agents That Reduce Work and Information Overload", Communications of the ACM, 1994, 37(7), pp. 31-40.

[14] J. McGrenere, and W. Ho, "Affordances: Clarifying and Evolving a Concept", Proceedings of Graphics Interface, May 2000, Montreal, Canada.

[15] M. Moundridou, and M. Virvou, "Evaluating the Persona Effect of an Interface Agent in a Tutoring System", Journal of Computer Assisted Learning, 2002, 18, 253-261. Blackwell Science.

[16] R. Moreno, R. E. Mayer, and J. C. Lester, "Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of Their Interaction", ED-MEDIA 2000 Proceedings, p.741-746. AACE Press.

[17] P. Murano, "Why Anthropomorphic User Interface Feedback Can be Effective and Preferred by Users", 7th Int.

Conference on Enterprise Information Systems, 2005, (c) – INSTICC, Miami.

[18] P. Murano, "Anthropomorphic vs Non-Anthropomorphic Software Interface Feedback for Online Factual Delivery", Proc. of the 7th. Int. Conference on Information Visualisation, 2003, IEEE, p. 138, London.

[19] P. Murano, "Anthropomorphic vs Non-Anthropomorphic Software Interface Feedback for Online Systems Usage", 7th ERCIM Int. Workshop on User Interfaces For All, 2002, pp. 339-349, Paris.

[20] P. Murano, "Effectiveness of Mapping Human-Oriented Information to Feedback From a Software Interface", Proc. 24th International Conference on Information Technology Interfaces, 2002, Cavtat, Croatia.

[21] P. Murano, "A New Software Agent 'Learning' Algorithm", People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, 2001, IEE, Manchester.

[22] P. Murano, "Mapping Human-Oriented Information to Software Agents for Online Systems Usage", People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, 2001, IEE, Manchester.

[23] P. Murano, C. Ede, and P. O. Holt, "Effectiveness and Preferences of Anthropomorphic User Interface Feedback in a PC Building Context and Cognitive Load", 10th International Conference on Enterprise Information Systems, Barcelona, Spain, 12-16 June 2008. (c) - INSTICC.

[24] P. Murano, and P. O. Holt, "Evaluation of an Anthropomorphic User Interface in a Telephone Bidding Context and Affordances", 12th International Conference on Enterprise Information Systems, Madeira, Portugal, 8-12 June, 2010, (c) - INSTICC. In Cooperation With ACM SIGCHI

[25] C. Nass, and S. Brave, Wired for Speech How Voice Activates and Advances the Human-Computer Relationship, 2005, The MIT Press.

[26] D. A. Norman, "Affordance, Conventions and Design". Interactions, May-June, 1999, p.39-42.

[27] D. A. Norman, The Design of Everyday Things, 2002, Basic Books.

[28] L. Qiu and I. Benbasat, "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems", Journal of Management Information Systems, Spring 2009, Vol. 25, No. 4, pp. 145-181, M.E. Sharpe, Inc.

[29] B. Shneiderman, and C. Plaisant, "Designing the User Interface: Strategies for Effective Human Computer Interaction", 2005, Pearson Education.

[30] A. Takeuchi, and T. Naito, "Situated Facial Displays: Towards Social Interaction". Proc. CHI'95 Human Factors in Computing Systems, 1995, pp. 450-454.

[31] M. Wooldridge, An Introduction to Multiagent Systems, 2009, 2nd Ed, John Wiley and Sons Ltd.

[32] J. Xiao, R. Catrambone, and J. Stasko, "Be Quiet? Evaluating Proactive and Reactive User Interface Assistants", Proceedings of INTERACT 2003, IOS Press, 383-390

[33] V. Zue, "Talking With Your Computer", Scientific American, 1999, 281, pp. 40-41,

Dr Pietro Murano is a Computer Scientist at the University of Salford, UK. Amongst other academic and professional qualifications he holds a PhD in Computer Science. His specific research areas are in Human Computer Interaction and Usability of software systems. He publishes regularly at refereed international levels and also supervises PhD level work at the University of Salford.

Patrik O'Brian Holt Is a Research Professor of Computing (Cognitive Engineering) at Robert Gordon University, Aberdeen. His research interests include interactive visualisation of complex data and cognitive modelling of human performance. Prof. Holt is the Director of the Joint Research Institute for Computational Systems which is part of the Northern Research Partnership in Engineering, a Scottish research pooling initiative.