EFFECTIVENESS AND PREFERENCES OF ANTHROPOMORPHIC USER INTERFACE FEEDBACK IN A PC BUILDING CONTEXT AND COGNITIVE LOAD

Pietro Murano, Christopher Ede
University of Salford, Computer Science, Multimedia and Telecommunications, Centre for Virtual Environments, Newton Building, Gt. Manchester, M5 4WT, UK
p.murano@salford.ac.uk, chris_ede9@yahoo.co.uk

Patrik O'Brian Holt
Interactive Systems Research Group, School of Computing, The Robert Gordon University, St. Andrew Street, Aberdeen, AB25 1HG, Scotland
p.holt@rgu.ac.uk

Keywords: Anthropomorphism, User interface feedback, Evaluation, Cognitive Load.

Abstract: This paper describes an experiment and its results concerning research that has been going on for a number of years in the area of anthropomorphic user interface feedback. The main aims of the research have been to examine the effectiveness and user satisfaction of anthropomorphic feedback in various domains. The results are of use to all interactive systems designers, particularly when dealing with issues of user interface feedback design. Currently the work in the area of anthropomorphic feedback does not have any global conclusions concerning its effectiveness and user satisfaction capabilities. This research is investigating finding a way for reaching some global conclusions concerning this type of feedback. The experiment detailed, concerns the specific software domain of software for in-depth learning in the specific context of PC building. Anthropomorphic feedback was compared against an equivalent non-anthropomorphic feedback. The results were not statistically significant to suggest one type of feedback was better than the other. It was also the aim to examine the types of feedback in relation to Cognitive Load Theory. The results suggest that the feedback types did not negatively affect Cognitive Load.

1 INTRODUCTION

The user interface is usually the most visible part of a software application and is often one of the main aspects that users really think about if they have a problem or if they like something about the application they are using. It is therefore very important to strive to achieve as good quality a user interface as possible as this will affect the users' perceptions about a system. In the worst case, if the user interface is unusable, the whole application could be abandoned.

The global aim of this research is to constantly discover better ways of developing user interfaces, specifically the feedback that is given to users. The research is directly concentrating on investigating the effectiveness and user satisfaction of anthropomorphic feedback. To achieve this direct comparisons are being made with non-anthropomorphic feedback in an experimental setting. Furthermore, the authors of this paper are also trying to explain the results of conducted experiments in terms of appropriate theories. One such theory that is being investigated in conjunction with the experimental results is the Cognitive Load Theory.

Anthropomorphism at the user interface usually involves some part of the user interface, taking on some human quality (De Angeli, Johnson, and Coventry, 2001). Some examples include a synthetic character acting as an assistant or a video clip of a human.

1.1 Background Literature

Other researchers have conducted some investigations into the area of anthropomorphic feedback, but the wide spread results of these efforts do not reveal an overall global picture indicating if such types of feedback are preferable (or not). One of the earliest studies into anthropomorphism at the user interface was by Quintanar, Crowell, Pryor and Adamopoulos (1982). This was an experiment in a quiz context which tested anthropomorphic textual feedback and non-anthropomorphic textual feedback. The quiz was about 'psychology' and the participant sample was undergraduate students. The main aims of the study were to obtain the user's thoughts about the system and ascertain the effectiveness of the two types of feedback. One finding of the study showed that participants perceived the anthropomorphic feedback to be 'more human, less honest and slightly less courteous…' compared with the non-anthropomorphic feedback. Also Quintanar et al (1982) found the anthropomorphic textual feedback to be more effective. This was based on the fact that the quiz scores were higher under the anthropomorphic condition and the amount of time spent thinking about the questions presented and the system's responses was higher. While this was an interesting study, it did not address other forms of anthropomorphism, such as synthetic characters or video clips etc. Also some methodological aspects could have caused some bias in the results. One such issue concerns the recruitment of the participants. The authors state that participants were screened for their psychology knowledge, which is appropriate. However they do not state how the different experimental conditions were balanced in terms of the participants' psychology skills.

Also, Moreno, Mayer and Lester (2000) have done some work in relation to anthropomorphic user interfaces. Two experiments described in (Moreno et al, 2000) looked at varying the kind of communication an agent used towards the user. With such variations they wanted to know if participants in a certain experimental group could have 'deeper understanding'. They also wanted to discover if deep understanding would be affected if using voice compared to text. The first study varied the agents in four different conditions in a botany context. The first condition used a synthetic type agent able to 'converse'. The second condition used only a 'conversing' agent (i.e. no image or animation). The third condition used a synthetic type agent communicating by means of text only. The fourth condition used only text. The second study used the same four conditions. However the synthetic agent and its corresponding synthetic 'voice' were replaced with a real human. Their results show that the presence of an on screen image of an agent did not significantly affect learning. However they obtained significant results in favour of using a voice to communicate information compared with text. Their results suggest that using a voice helps to improve learning. They found experimental participants' ability to remember and problem solving skills to be improved. Also the voice agent was rated more highly by experimental participants. However some experimental design flaws suggest that more work is required in this area. Their publication (Moreno et al, 2000) does not detail clearly enough if the participants were screened properly for their prior botany knowledge. Some screening did take place, but how this was done is not detailed and could have biased the results if some participants, e.g. happened to have more botany knowledge in one condition. Also the paper does not detail the level of difficulty of the botany material used in the context of the experiment in relation to the participants' experience.

A further study concerning tutoring by Moundridou and Virvou (2002) tested 2 conditions in an algebra tutoring environment. The participants were screened in advance for their mathematical knowledge by means of a test and were deemed to be approximately equivalent to one another. The first experimental 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. Lastly the participants were given a post-experiment test and this did not reveal any statistically significant difference between the conditions in terms of overall results.

Further, the authors of this paper have been investigating the appropriate use of anthropomorphic feedback for some time and the results obtained are not definitive with respect to obtaining global knowledge in this area, e.g. (see Murano, 2005, 2003, 2002a, 2002b, 2001a and 2001b). However related to the experiment reported in this paper, the study by Murano (2002b) in the same domain of online learning and the specific context of English pronunciation, showed with significant results that using an anthropomorphic feedback was more effective and preferred by users. This in effect meant that the anthropomorphic condition aided the correction process more. This was an experiment which used Italian participants with imperfect English pronunciation. Several tasks involving pronunciation exercises were used where either an anthropomorphic or non-anthropomorphic feedback was used to assist in the correction process. The anthropomorphic feedback consisted of a video of a human and the non-anthropomorphic feedback consisted of guiding text and a diagram.

2 PC BUILDING EXPERIMENT

2.1 Aims and Objectives

Therefore this paper investigates further the domain of online learning. This time it is in the specific context of PC building. The aim of this experiment was to gather data regarding effectiveness and user satisfaction in the PC building context. Specifically the aim was to find out if anthropomorphic user interface feedback fostered a better interaction experience with fewer errors and therefore a better task completion rate. It was also of interest to find out if anthropomorphic user interface feedback led to better user satisfaction.

The authors were also interested to find out if cognitive load was a factor in some of the results observed. This theory essentially argues that when the overall cognitive load exceeds a particular threshold, then activities such as learning are impaired or become more difficult (Martin-Michiellot and Mendelsohn, 2000). Furthermore cognitive load has three basic strands. The first is 'intrinsic cognitive load'. This concerns the activity involved in learning some item of information and typically how many units of information are being learned at the same time. The second strand concerns 'extraneous cognitive load'. This has to do with how learning materials are presented to a human. Therefore the more complex the manner of presentation is, the higher the extraneous cognitive load tends to be (Martin-Michiellot and Mendelsohn, 2000). The third strand concerns 'germane cognitive load'. This strand has to do with the human faculty of processing and understanding information, and problem solving (Sweller, van Merrienboer and Paas, 1998).

2.2 Users

- All the participants taking part in the study were of varied age groups.
- 30 participants were used in the experiment. Although gender was not the main aspect under consideration the sample used was approximately 50:50 for gender.
- The participants had varied occupations.
- All the participants were novices to the area of PC assembly. This ascertained by administering a small pre-experiment test and only those with low scores, (i.e. little knowledge about hardware etc) were used.

2.3 Experimental Design

A between users design was used. The 30 participants were randomly assigned to one of the 2 conditions being tested –anthropomorphic or non-anthropomorphic. Randomness was achieved by alternately assigning a participant to one of the conditions until all 30 had been assigned.

2.4 Variables

The independent variables were the types of feedback, i.e.:

- Textual instructions.
- MS Agent synthetic character.

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 counting the number of errors made as each participant attempted to assemble the components, whether the participant completed a task and the time taken to complete a task. These factors were then used in a scoring formula in order to achieve a single score per participant (see note below). The formula was devised so as to allow groups of similar types of errors to be catalogued under a dedicated category. The errors were determined by the experimenter physically examining the way the components had been inserted into the PC case (see note below). The time taken to complete the tasks and the number of times a participant clicked a certain button in the application were recoded automatically by the application. The subjective opinions were measured by means of a post-experiment questionnaire (Appendix 23).

(NOTE – The formula used was as follows:

- Each participant (unknown to them) was started on 5 points for each task.
- For a completed task with 1 minor error, 1 point was deducted. A minor error was of the kind that led to a device working, but still having some problem, e.g. not securing the CD ROM drive to the case with screws.
- For a completed task with 1 major error, 2 points were deducted. A major error was of the kind leading to a device not working but essentially still being fitted in place, e.g. fitting the CD ROM drive appropriately but not being able to insert the required power cable.
- For a completed task with 1 severe error, 3 points were deducted. A severe error was 2 or more major and minor errors e.g. not securing the CD ROM drive to the case with screws was a minor error and not being able to insert the required power cable was a major error.
- For an unsuccessful attempt at completing a task, 4 points were deducted, e.g. a participant tried to do a task, but then gave up or a participant tried to do a task, but made 2 or more severe errors.
- If no attempt at all was made for a task, all 5 points were deducted to leave a score of 0.)

2.5 Apparatus and Materials

- A PC running Windows XP with 256 Mb RAM.
- Microsoft Agent 2.0 ActiveX component. Lernout and Hauspie TruVoice Text-To-Speech engine. The prototype was engineered with VB.NET.
- External speakers.
- Desktop microphone used for the anthropomorphic condition.
- An open PC case with the motherboard, processor and power supply already assembled.
- Disassembled RAM board, CD ROM drive, ribbon cable and relevant assembly screws.
- Screw driver.

2.6 Procedure and Tasks

The first step was to recruit a suitable number of participants particularly meeting the requirement of being novices to the assembly of computer hardware. As stated above this was achieved by the participants completing a pre-experiment questionnaire and a small pre-experiment test covering basic knowledge of PC components (only participants with a low score in the pre-experiment test were used – a high score would have indicated too much knowledge for the experiment). Each participant was briefed with the following points before commencing the actual experiment:

- The software was developed for evaluating its suitability to teaching about PC building.
- Help could be received from the system by using the help button (non-anthropomorphic group) or by asking for help via the microphone (anthropomorphic group).
- The information for each stage would be shown for a limited period of time. If the information was to be repeated, this could be achieved by pressing the 'repeat' button (non-anthropomorphic group) or by asking for the information to be repeated via the microphone (anthropomorphic group).
- Video demonstrations were available for viewing concerning the assembly of each part.
- It would not be possible to backtrack to a previous stage on screen.
- A post-experiment questionnaire would need to be completed at the end of the experiment.

Then the procedure described below was carried out in the same way for all participants using the same environment, equipment and questionnaires/observation protocols. Each participant was treated in the same manner. This was all in an effort to control any confounding variables.

There were 2 tasks involving PC building. Specifically, the first task concerned inserting a RAM board into its appropriate slot in the motherboard. The second task involved assembling a CD ROM drive into the case, with a correct master jumper setting and connecting the necessary cables and screws.

Each participant was booked an appointment during the day. The experiment took about 30 minutes to complete per volunteer. After completing the initial pre-experiment questionnaire and test, participants were able to view the PC case and motherboard information screens. This allowed the participant to become familiar with the software before taking part in the assembly tasks.

Then for the first task, the participant was directed to press the 'Random Access Memory Walkthrough button'. This would then initiate a series of steps, with accompanying photographs of the parts etc, describing what had to be done to complete the task. For the anthropomorphic condition, the Merlin character would narrate the various aspects required and where relevant would move on the screen and 'point' to certain elements on the photographs. For the non-anthropomorphic condition the same content was delivered textually and to match the Merlin character's movement and pointing to various elements on the photographs, clearly visible arrows were used to 'point' to the same elements on the screen. The textual information was displayed on the screen for the same amount of time as the Merlin character took to narrate the information. However the participants in each condition had the option of having the information repeated whilst within a particular information stage.

Once the information had been received the participant would attempt to physically insert the RAM into the appropriate slot. Although the actual information regarding the RAM board could not be viewed again at this stage, the participant did have the software's help available. Also the information regarding the PC case and motherboard which had been made available before starting the first task of inserting the RAM was available at this stage. Lastly the relevant video demonstration was playable by the participant at this stage. Once the first task was completed, the second task was undertaken in the same manner as described above.

At all times during the experiment the participants were informally observed and at the end of a task, an inspection of the relevant components and their positioning etc. was carried out by the experimenter – any errors being noted.

Lastly the participants were asked to complete a post-experiment questionnaire regarding their subjective opinions of their experience with the software.

2.7 Results

The data collected was analysed using MANOVA analysis. Regarding effectiveness, the amount of time taken for each task was recorded, the number of errors committed were recorded and also if a task was

completed was recorded. For user satisfaction, various subjective opinions were elicited from the participants.

As discussed in section 2.4 above, the factors of task completion and errors were used in a scoring formula. The results for the first task involving the RAM board are presented below in Table 2.1. The analysis involved the actual scores obtained, the experimental group, the age groups of the participants and the gender of the participants.

Table 2.1 Analysis of Variance, Task 1, Overall Score

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	5.161467	1.03229	1.1362
Error	24	21.805200	0.90855	Prob > F
C. Total	29	26.966667		0.3685

Table 2.2 Analysis of Variance, Task 2, Overall Score.

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	10.994271	2.19885	1.8008
Error	24	29.305729	1.22107	Prob > F
C. Total	29	40.300000		0.1508

Table 2.3 Analysis of Variance, Task 1, Times.

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	67156.03	13431.2	2.4924
Error	24	129330.63	5388.8	Prob > F
C. Total	29	196486.67		0.0593

Table 2.4 LSMeans Differences Student's t
Alpha = 0.050 t = 2.0639LSMean[i] By LSMean[j].

Mean[i]-Mean[j] Std Err Dif Lower CL Dif Upper CL Dif	female	male
female	0 0 0 0	82.7179 26.9818 27.0302 138.406
male	-82.718 26.9818 -138.41 -27.03	0 0 0 0

Level			Least Sq Mean
female	A		387.47313
male		B	304.75522

Levels not connected by same letter are significantly different

Table 2.5 Analysis of Variance, Task 2, Times.

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	54857.62	10971.5	0.6896
Error	24	381815.34	15909.0	Prob > F
C. Total	29	436672.97		0.6361

As with the previous task, the F-ratio of 1.80 shows that there are no significant differences between the various factors analysed.

The F-ratio of 1.14 shows that there are no significant differences between the various factors analysed.

Also the same analysis was carried out for the second task involving the fitting of a CD ROM drive. This is shown in Table 2.2 below. The analysis involved the actual scores obtained, the experimental group, the age groups of the participants and the gender of the participants.

The tasks were also timed and included in the analysis. For the first task (RAM Board insertion) the analysis involved the actual times obtained, the experimental group, the age groups of the participants

and the gender of the participants. This is shown in table 2.3 above.

The F-ratio of 2.49* is tending towards significance ($p < 0.05$). The initial results were then subjected to post-hoc testing using a t-test, where significance was shown for the gender, where female participants were significantly slower than the male participants. This can be seen in Table 2.4 above.

For the second task (CD ROM fitting) the analysis also involved the actual times obtained, the experimental group, age group and gender. This is shown in table 2.6 above.

Strictly the F-ratio of 2.12 is only approaching significance ($p < 0.05$). However the initial results were then subjected to post-hoc testing using a t-test, where significance was shown for the experimental group, where the anthropomorphic group significantly considered the application more helpful than their counterparts in the non-anthropomorphic group. This can be seen in Table 2.7 above.

A further aspect of interest concerns the application being frustrating to use. This was in relation to the experimental group, age group and gender. This is shown in table 2.8 above.

group, the age groups of the participants and the gender of the participants. This is shown in table 2.5 above.

The F-ratio for the task involving the times shows no significance for the various factors being analysed.

Regarding the participant subjective responses various aspects were considered. The main aspect of interest concerns the application being considered helpful for PC assembly. This was in relation to the

Table 2.6 Analysis of Variance, Application Considered Helpful.

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	6.979297	1.39586	2.1175
Error	24	15.820703	0.65920	Prob > F
C. Total	29	22.800000		0.0981

Table 2.7 LSMeans Differences Student's t
Alpha = 0.050, t = 2.0639LSMean[i] By LSMean[j].

Mean[i]-Mean[j] Std Err Dif Lower CL Dif Upper CL Dif	Anthropomorphic	Non-Anthropomorphic
Anthropomorphic	0 0 0 0	0.81598 0.32733 0.1404 1.49157
Non-Anthropomorphic	-0.816 0.32733 -1.4916 -0.1404	0 0 0 0

Level			Least Sq Mean
Anthropmorphic	A		7.7819451
Non-Anthropomorphic		B	6.9659605

Levels not connected by same letter are significantly different

Table 2.8 Analysis of Variance, Application Frustrating to Use.

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	18.272845	3.65457	3.3699
Error	24	26.027155	1.08446	Prob > F
C. Total	29	44.300000		0.0191

This result is significant ($p < 0.05$) with an F-ratio of 3.37*, particularly in relation to the experimental groups. The non-anthropomorphic group significantly rated the application as more frustrating to use compared to their counterparts in the other anthropomorphic condition.

2.8 Conclusions

As can be seen in the previous section various aspects were statistically analysed. Regarding the effectiveness issues, no statistical significance can be seen in the scores of the two groups which included the task completion successes and the errors made by participants.

The times were also analysed and for the first task, there was no statistical significance for the experimental group. However there was statistical significance to show that the female participants were slower than the male participants. Although this research is not primarily about gender, it would have been interesting to know why this result is in this direction. Unfortunately none of the female (or any of the participants) participants were available again for interview. Also the initial demographic data that was collected as part of the recruitment process does not reveal any information that could enlighten the authors on the matter. However the authors believe that it is possible that something in the female group's background could shed light on the matter. Further, for the second task, there is no statistical significance for the time taken to complete the task. This is for the experimental group and the gender.

The user preference issues analysed suggest that overall the preferences tended towards the anthropomorphic condition. Participants in the anthropomorphic condition rated the application as more helpful than the participants in the non-anthropomorphic condition. Also the non-anthropomorphic group significantly rated the application as more frustrating to use than the anthropomorphic condition participants. Despite the statistical analysis indicating a user preference towards the anthropomorphic feedback, the authors suggest caution in categorically declaring the anthropomorphic feedback as being more satisfying to use. This is because as described in section 2.6 above, the non-anthropomorphic condition had the feature incorporated where the textual information remained on the screen only for a certain amount of time. This aspect was designed into the system to more closely match the fact that the anthropomorphic feedback condition, consisting of a character with accompanying speech bubbles, only appeared on the screen for the time it took the character to utter the information. In essence the authors did not want the non-anthropomorphic condition participants having an unfair advantage by having the information available to them for a much longer period of time. While it is argued that this feature should have balanced the two conditions more closely, the authors suggest that this potentially incurred the side effect of participants in the non-anthropomorphic group rating the helpfulness of the application significantly lower than the other participants in the other experimental group and rating the application as significantly more frustrating to use.

Comparing the results for this experiment with the results of previous work carried out by Murano (2002b), it is clear that the effectiveness issues do not match because in the English pronunciation context the anthropomorphic feedback was more effective, while in the experiment detailed above there were no significant differences in the two tested conditions. The only slight agreement between the two experiments concerns the subjective opinions of the participants. In both cases it appears that participants tend to prefer the anthropomorphic feedback given the similar domain and contexts being considered. However as argued in the previous paragraph, the subjective responses for the experiment described above need to be tempered with the knowledge that the attempts at having well balanced experimental conditions may have caused more negative opinions being given towards the non-anthropomorphic feedback.

3 THE EXPERIMENT AND COGNITIVE LOAD

As stated in the introduction of the paper, it was also of interest to know if cognitive load could have been a factor somewhere in the conditions and user interface being tested. Informally, the fact that there were no significant differences in effectiveness could indicate that the cognitive load was equivalent under both conditions. However the following discussion in relation to cognitive load should help to make things clear.

Firstly the participants were definitely beginners to PC assembly as indicated by the recruitment process. This characteristic could mean there could have been some intrinsic cognitive load as they were all learning something new to them with a few units of information. However this would have spanned both conditions and not been affected by the feedbacks being tested, so therefore should have been the same under both conditions. Regarding the possibility of extraneous cognitive load, the very nature of the character moving on the screen could have increased this form of cognitive load, as perhaps more integration on the part of the user would have been needed, compared with the static textual explanations which were definitely in integrated format. The possibility of one condition incurring a higher germane cognitive load is unlikely, as the content of the information was the same in both conditions. Also one could argue that the textual information was static and perhaps gave more advantage to the user compared to the Merlin character's speech and speech bubbles. However this was not the case as the textual information was not on the screen all of the time as discussed above. Therefore if cognitive load issues had been at play, it would have been expected that participants in the non-anthropomorphic condition should have overall had significantly more success in the task – this was not the case though. Lastly this experiment timed the participants whilst they did the tasks and generally no significant result was observed between the conditions (except for female participants). The reason this is also of interest is that cognitive load has been linked to task completion times as an indicator of increased cognitive load (Neerincx, van den Dobbelsteen, Grootjen and van

Veenendaal, 2003 and Wang, Kaufman, Mendoca, Seol, Johnson and Cimino, 2002).

It can therefore be argued that the two conditions being tested were approximately equivalent in terms of cognitive load. As stated, if cognitive load had been a significant factor in one of the feedbacks, it would have been expected that the performance of participants in the anthropomorphic condition should have had worse scores and worse overall times for task completion.

3.1 Overall Conclusions

As can be seen by the results of this experiment and also the previous work briefly considered in this paper, there is still work to be done in this area to try and determine the suitability of such types of feedback. More work is being carried out. Results are being further statistically analysed and also analysed in terms of other theories of cognition and human processing.

ACKNOWLEDGEMENTS

The authors would like to thank for their support: The School of Computing Science and Engineering, University of Salford. Heriot-Watt University, Edinburgh, Dept. of Computer Science. The Robert Gordon University, Aberdeen, School of Computing.

REFERENCES

De Angeli, A, Johnson, G. I. and Coventry, L., 2001. The Unfriendly User: Exploring Social Reactions to Chatterbots, *Proceedings of the International Conference on Affective Human Factors Design*, Asean Academic Press.

Martin-Michiellot, S. and Mendelsohn, P., 2000. Cognitive Load While Learning With a Graphical Computer Interface, Journal of Computer Assisted Learning, 16, 284-293, Blackwell Science.

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

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

Murano, P., 2005. Why Anthropomorphic User Interface Feedback Can be Effective and Preferred by Users, *7th International Conference on enterprise Information Systems*, Miami, USA, 25-28 May 2005. (c) - INSTICC.

Murano, P., 2003. Anthropomorphic Vs Non-Anthropomorphic Software Interface Feedback for Online Factual Delivery *7th International Conference on Information Visualisation (IV 2003) An International Conference on Computer Visualisation and Graphics Applications*, London, England, 16-18 July 2003, (c) - IEEE.

Murano, P., 2002a. Anthropomorphic Vs Non-Anthropomorphic Software Interface Feedback for Online Systems Usage, *7th European Research Consortium for Informatics and Mathematics (ERCIM) Workshop - 'User Interfaces for All' - Special Theme: 'Universal Access'.Paris*(Chantilly),France 24,25 October 2002. Published in Lecture Notes in Computer Science (C) - Springer.

Murano, P., 2002b. Effectiveness of Mapping Human-Oriented Information to Feedback From a Software Interface, *Proceedings of the 24th International Conference on Information Technology Interfaces*, Cavtat, Croatia, 24-27 June 2002.

Murano, P., 2001a. A New Software Agent 'Learning' Algorithm, *People in Control An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*, UMIST, UK, IEE.

Murano, P., 2001b. 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*, UMIST, UK, IEE.

Neerincx, M. A., van den Dobbelsteen, G. J. H., Grootjen, M., van Veenendaal, J., 2003. Assessing Cognitive Load Distributions for Envisioned Task Allocations and Support Functions, *Thirteenth International Ship Control Systems Symposium*, Orlando, Florida, USA, 7-9.4.03.

Quintanar, L. R, Crowell, C. R, Pryor, J. B. and Adamopoulos, J., 1982. Human Computer Interaction: A Preliminary Social Psychological Analysis, Behaviour Research Methods and Instrumentation 14(2): 210-220

Sweller, J., van Merrienboer, J. J. G. and Paas, F. G. W. C., 1998. Cognitive Architecture and Instructional Design, Educational Psychology Review, 10(3), Plenum.

Wang, D., Kaufman, D. R., Mendonca, E. A., Seol, Y., Johnson, S. B., and Cimino, J. J., 2002. The Cognitive Demands of an Innovative Query User Interface, *Proceedings of the AIMA 202 Annual Symposium*, 850-854.