An Empirical Study on the Effects of Embodied Conversational Agents on User Retention Performance and Perception in a Simulated Mobile Environment

Ioannis Doumanis[,1] and Serengul Smith
Middlesex University, The Burroughs Hendon, London NW4 4BT

Abstract. The paper presents a user study designed to examine the impact of the presence of a multimodal ECA on the user's ability to retain content of cultural value with variable degree of difficulty (i.e., technical and simple content). The study was conducted in the lab, using a high resolution panorama representing four locations in an archaeological attraction. The content participants perceived differed both in terms of complexity and length. Participants interacted with an ECA-based system and then with a non-ECA system that provided content about popular locations in the attraction. Results indicate that participants who used the system with the ECA retained content of variable difficulty more consistently, than those who used the system without the ECA. However, we also found that if text is added as an additional output modality to an ECA-based information system it can positively impact the perception of the technical content, which can potentially lead to enhanced retention of technical content.

Keywords: Embodied conversational agents, human-centered computing, mobile tour guides, information systems

Introduction

Evolution in the area of mobile computing has been phenomenal in the last few years. The exploding increase in hardware power has enabled multimodal mobile interfaces to be developed. These interfaces differ from the traditional graphical user interface (GUI), in that they enable a more "natural" communication with mobile devices, through the use of multiple communication channels (e.g., multi-touch, speech recognition, etc.). As a result, a new generation of applications has emerged that provide human-like assistance in the user interface (e.g., the Siri conversational assistant) [1]. These conversational agents are currently designed to automate a number of tedious mobile tasks (e.g., to call a taxi), but the possible applications are endless. A domain of particular interest is that of Cultural Heritage, where conversational agents can act as personalized tour guides in, for example, archaeological attractions. The visitors to historical places have a diverse range of information needs. For example, casual visitors have different information needs from those with a deeper interest in an

---

[1] Corresponding Author: *Middlesex University, The Burroughs Hendon, London NW4 4BT;* Email: ioannis.doumanis@lsst.ac

attraction (e.g., - holiday learners versus students). A personalized conversational agent can access a cultural heritage database, and effectively translate data into a natural language form that is adapted to the visitor's personal needs and interests. The study presented in this paper focuses on the information needs of a specific type of visitors, those for whom retention of cultural content is important (e.g., students of history, cultural experts, history hobbyists, educators, etc.). Embodying a conversational agent enables the agent to use additional modalities to communicate this content (e.g., through facial expressions, deictic gestures, etc.) to the user. Simulating the social norms that guide the real-world human-to-human interaction (e.g., adapting the story based on the reactions of the users), should at least theoretically optimize the cognitive accessibility of the content. Although a number of projects have attempted to build embodied conversational agents (ECAs) for cultural heritage [2, 3], little is known about their impact on the users' perceived cognitive accessibility of the cultural heritage content, and the usability of the interfaces they support. In particular, there is a general disagreement on the advantages of multimodal ECAs in terms of users' task performance and satisfaction over non-anthropomorphised interfaces.

We present a user study designed to evaluate the impact of the presence of a multimodal ECA on the cognitive accessibility of a mobile tour guide system, providing cultural content of variable difficulty. The system consists of a no-ECA version (control condition) and a version with an ECA capable of augmenting the acoustic information with relevant non-verbal behaviours (e.g., pointing gestures to direct user's attention focus to aspects of interest of a particular location). In a within-subjects design 14 participants interacted with both systems each with cultural content of variable complexity (i.e., simple and technical content). We evaluated the question of the impact of the presence of the ECA on the retention performance and perceived cognitive workload, as an indication of how participants felt about using the systems to uncover information about the specified locations. The study was conducted in the lab using a high-resolution panoramic application (see Figure 1).

The panoramic application included the photographic representation of four locations of the archaeological attraction and there were no on-screen buttons for users to interact with. An on-screen menu (see Figure 1), allowed users to visit the locations in any order they liked. To start the information presentation the system had to decode the name of the location from a QR-Code (shown as a camera icon in Figure 1) embedded in each location the user visited. A QR-Code is a simple two-dimensional bar-code that can be used as a cheap solution for physical location/object identification. In the current implementation the user has to photograph a QR-Code using the integrated camera of the tablet device. We had the following hypotheses:

- *"The degraded retention of complex information"*- The presence of a multimodal ECA has a negative impact on the retention of the technical information, for example, because it adds an extra burden to the already overloaded cognitive resources of the user (because of the complex nature of the technical information), but neither a positive nor a negative effective for the simple information.
- *"The enhanced retention of complex information"*- The presence of a multimodal ECA increases the participant's retention performance with the technical information, for instance because it reduces the cognitive loads (e.g., by rendering the interaction smoother) needed for retaining such information, but has no effect (neither positive nor negative) on the simple information.
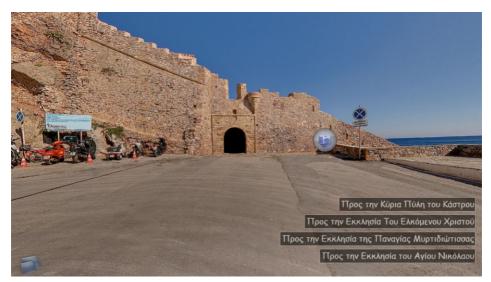
Figure 1. A screenshot of interactive panoramic applications

1. Related and Previous Work

The impact of Embodied Conversational Agents (ECAs) on the user's ability to retain information has been analysed in a number of prior studies [2, 3]. In one study [2], the impact of a mobile effective tour guide was evaluated on the users' recall performance under realistic mobile conditions. Three different types of mobile guide were evaluated in an interactive tour of the "Los Alamos" site of the Manhattan project. All stories generated by the system, related to the "Making of the atomic bomb". The physical tour, however, took place at the Heriot-Watt Edinburgh campus buildings, where buildings from the "Los Alamos" site were mapped onto University buildings. The agents differed in terms of emotions and attitude, portrayed by a simplistic 2D cartoon-like head, and by the inclusion of the agent's perspective and experiences in the narration. The fully affective guide could exhibit both emotions and attitude while the other two displayed no emotions nor attitude, or emotions but no attitude respectively. For example, the emotional guides could dynamically update the story to include their own feelings and perspective about historical facts (e.g., *"it seemed brutal to be talking about burning homes"*). The participants were requested to listen to at least three stories at each location, under a thematic area of their choice (e.g., Science or Military). After the completion of each story, participants had to rate the degree of interest of the stories, as well as how much they agreed with the guide's argument. In the first group (i.e., the fully affective guide), the input given influenced the processing conducted by the guide, while in the other two it merely gave the impression that it did. Upon completion of the tour, participants had to answer two sets of questionnaires, one to indicate their subjective experience of the system, and another to test their recall levels of the information they listened during the tour. In terms of recall performance, the researchers found no significant differences in the users' recall levels of the presented information between the three guides. They attributed this to various confounding variables, such as the speed of the guide's voice, non-native English speaking users, etc.

We find this non-effect result to be rather expected. Although the conclusion that a guide with attitude and intelligence makes the interaction more interesting may be a valid conclusion, it is how these behavioural attributes are portrayed through non-verbal means that can translate subjective views into enhanced retention performance. For example, studies have shown that a more realistic depiction of a virtual human [4] can create greater participant involvement in a virtual experience. In addition, the absence of body language (e.g., beat gestures) deprived the guide of a valuable communication channel that would augment the presented information and, in turn, lead to greater user retention performance. Last and perhaps most importantly, the pretended Los Alamos site at the University campus made it impossible for the users to connect the content of the stories with their surrounding environment, thus causing unnecessary cognitive overload [5]. If the user's cognitive resources were devoted into blocking the external stimuli in order to focus his/her attention to the stories, it is not surprising that the attitude and emotions of the agent had no impact on his/her overall retention performance. If the study had been conducted on the actual site, where users would have the ability to physically visit the various sites, the results may have been different.

In another study [3], the impact of an ECA on motivation and learning performance in a repeated task over a period of time was evaluated in a laboratory setting. In the study, each group of participants experienced a vocabulary trainer application, either with an ECA (*with-agent* version) or without an ECA (*no-agent* version). In the no-agent version, the user interface consists of two windows displaying the English and German expressions and a row of buttons for showing and rating the answer. In the with-agent version, a female ECA was added in the middle of the screen featuring some idle movements to make her look alive and with a minimum amount of gestures. The researchers found a "Persona Zero-effect", i.e., that they found neither positive nor negative effects on motivation and learning performance. Therefore, it was concluded that adding an ECA on an interface does not benefit performance but also does not distract. The no-effect results produced by this study, is in fact, encouraging. If the mere presence of an ECA (with minimal or without nonverbal communicative behaviours) in a learning environment has no detrimental effects on performance, then it could be assumed that endowing ECAs with a full repertoire of proper nonverbal behaviours might, in fact, improve performance. Comprehension can be directly affected by redundancy [6] and, hence, the use of an additional, redundant channel of communication, such as gestures or facial expressions, could result in more learning. In particular, the use of gestures could reduce message ambiguity by focusing learner attention, and facial expression can reflect and emphasize the agent message, emotions, personality and other behaviour variables [7]. Our study attempts to examine the validity of this hypothesis.

2.   Prototype systems

For the current study, we developed two tour guide applications one featuring a multimodal ECA and another without an ECA on the interface. Each system, used photographs of each of the locations the user would encounter as a background. The ECA could refer to objects in its background and to additional information that appeared in a floating 3D window. A dialogue window provided users with the ability to have a short "get-to-know-each-other" dialogue with the system (e.g., about how to

use the system). A control window provided access to the device's on-board camera and to an interactive map of the castle that showed the locations users had to visit. The non-ECA system (see right side of Figure 2) features the same interface elements, but instead of an ECA a subtitle window "reads" the system contents while highlighting each word of the text.

Figure 2. The system with the ECA (left side) and the system without the ECA (right side)

3. User Study

Population: In total, fourteen users (both males and females) from a variety of age groups took part in this study. The participants were randomly assigned to two groups of seven. None of the participants was either a local-resident or had visited the area before. This was done to avoid over-familiarity with the area. All participants were native Greek speakers and had a variety of academic and mobile-computer backgrounds.

Task: The goal of this experiment was to investigate different information presentation systems capable of providing content of variable difficulty about attractions in the castle, with respect to their effects on their ability to effectively retain information. Participants were asked to use an interactive panoramic application to visit four locations in the castle (in any order they like) and retrieve information with varying degree of difficulty (simple then technical or vice versa), once using system A (i.e., with the ECA) and once using the system B (i.e., without the ECA). The technical content was a technical description of the locations, while the simple content was taken from the information leaflet the castle provides for free to all visitors. The total duration of each tour was not more than 20 minutes. Furthermore, participants were informed that an experimenter would be present in the lab to observe their behaviour while using the system and to provide help if necessary (e.g., if they could not use the camera to photograph a QR-Code). In addition, they were told that after visiting all locations, they would be asked to indicate in a test what they retained from the presentations. At the beginning of each task, the system asked participants to provide their personal details (i.e., name, gender and age) and to parameterise various features of the agent and the system (e.g., the ECA's appearance, volume, etc.). After that, a computer agent appeared either in the form of an ECA or a disembodied voice with a subtitle window. In order to start a presentation, participants had to click on a button embedded in each of the locations they visited using the panoramic applications. The button activated a QR-Code that users had to photograph using the device's camera. Once the QR-Code was decoded, the system would present the relevant information

about the particular location (simple or technical). After completely uncovering information for all four locations, participants were asked to indicate, on a five-point scale, whether they found the presentations difficult. Next, a retention test was administered which asked questions about the information they heard in each of the locations. Finally, participants were asked to indicate, their perceived cognitive workloads associated with the presentations they experienced with each of the systems, on a seven-point scale questionnaire.

Measures and Methods: The only objective variable in the experiment was the answers to the retention test. The subjective measures were the responses to the items of the questionnaire, and the ratings of the difficulty of the presentations. The retention test used the same fill-in-the-blanks approach. The questionnaire items used the same seven-point agree-disagree Likert [8] format (1=strongly disagree, 7=strongly agree), and measured the perceived cognitive workload, as an indication of how the participants felt about using the systems to uncover information about the specified locations.

4.   Results and Discussion

Performance Measures: We measured the amount information participants recalled from each type of content as an indicator of the effectiveness of each system (ECA present or ECA absent) in eliciting recall performance. A series of 2 x 2 ANOVAs, taking the score and confidence as dependent variables, and type of ECA (ECA-present vs. ECA-absent), order of presentation (simple then technical vs. vice versa) and type of content (simple vs. technical) as independent variables did not show any significant effects of any of the independent variables. There were no significant interactions either.

Table 1. Mean retention performances.

Order of presentation	ECA (Content) (n = 14)	Mean	Std. Deviation
Simple / Technical	Present (Simple)	25.1	13.4
	Absent (Technical)	21.8	10.99
Technical / Simple	Present (Technical)	21.1	11.9
	Absent (Simple)	36.8	22.4

However, (see Table 1) the participants' performances were more consistent with content of varying difficulty with the system with the ECA, than with the system without the ECA. Participants using the system with the ECA performed almost the same between the two content conditions (mean S/T = 25.1 vs. mean T/S = 21.1). Those participants that used the system without the ECA performed better with the simple content (mean Simple = 36.8) than with the technical content (mean Technical = 21.8). This is a strong indication that the modalities used by the ECA (voice, gestures, etc.) were more effective in enhancing the participants' ability to retain information of variable difficulty about the locations of the castle than the modalities used in the system without the ECA. This finding invalidates the hypotheses examined (*"The degraded retention of complex information"* and *"The enhanced retention of complex information"*) as the ECA does not result in enhanced or degraded retention

performances. In fact, it has no measurable impact on either the simple or the technical information content. Conversely, the variation of the content affected the participants using the system without the ECA. Their performance was better with the simple content than with the technical content. Hence, we can safely say that the presence of an ECA does not enhance information retention, but it can provide a more consistent method of presentation for cultural content of variable difficulty, than a system without such an artefact on the interface.

Subjective Assessment - Workload Questionnaire: Participant subjective impression of the cognitive workload required to navigate the routes was measured by nine-sets of questionnaire items rated on 7-point Likert scales. The first set of items assessed the complexity, learnability, consistency and self-organization requirements of the information task, and it includes six questions (Cronbach's α = 0.850). The second set includes four items, and it was designed to assess how participants perceived the output modalities (visual, auditory and textual) of the prototypes in terms of sensory, satisfaction, and understanding (e.g. visibility of the screen, confusion caused by the multiple modalities, etc.) (Cronbach's α = 0.646). The third set evaluated how the participants perceived the feedback they received from the prototypes in terms of sensory, timing, relevance and memory requirements (e.g., relevance of the output to the environment of the castle, support to photograph correctly the QR-Codes, etc.). This question set includes six items (Cronbach's α = 0.653). The fourth question set evaluated the working memory requirements of the prototypes. It includes four items (e.g., amount of information to hold in mind when using the prototypes, how the system should respond when a participant is confused\overloaded with information, etc.) (Cronbach's α = -0.029). The fifth set assessed the emotional impact of the prototypes (e.g., frustrating, annoying, etc.), and it includes four items (Cronbach's α = -4.410). The sixth set assessed how the prototypes impact the participants' long-term memory in terms of the task learnability and in relation to their existing knowledge (e.g., ease of learning of the information, relation of information to the participant's interests, etc.) (Cronbach's α = 0.503). This question set consists of four items. The seventh set evaluated how effectively participants could access the underlying structure of the information task, and it consists of four items (e.g., simplicity of the presented information, how the structure of the information is presented, etc.) (Cronbach's α = 0.700). The eighth set assessed how the participants perceived the rationality of their responses, and how supported they felt during their responses. This question set consists of four items (e.g., allowances for response errors, frequency of response errors i.e., wrongly retained information, etc.) (Cronbach's α = 0.446). The final set evaluated how the participants perceived their output responses and how supported they felt in order to respond appropriately. It includes four items (e.g., ease of finding the selected locations, support to learn the information provided, etc.) (Cronbach's α = 0.536).

For all analyses, we performed a series of 2 x 2 ANOVAs, taking each questionnaire item as a dependent variable, the type of ECA, the type of content, and order of presentation as independent variables. We found significant effects and interactions on the following items:

- "The information task is too complex"
- "The process of extracting information from the system is difficult to learn"

- "It's hard to learn any of the information presented by the system"
- "The completion of the information task requires too much self-organization"

On the difficulty of the information task there was an effect of the order of presentation (F (1, 24) = 9.422; p < .01) and an interaction between the type of content and the type of ECA (F (1, 24) = 9.422; p < .01). The significant interaction between the type of ECA and the content was further analysed using simple main effect analysis. It revealed that the variation of content across the order conditions significantly influenced how the participants perceived the complexity of the information task with both the system with the ECA (F (1, 24) = 4.342; p < .05) and the system without the ECA (F (1, 24) = 5.095; p < .05). Participants perceived the technical content as less complex than the simple content, when each type of content was presented by the system without the ECA. Then, they perceived the simple content as less complex than the technical content when each type of content was presented by the system with the ECA. The technical content may have seemed easier with the system without the ECA because of the text and voice used as output modalities. Then, because the presence of an ECA makes an interface more user friendly and the simple content was simple enough to understand with or without the ECA, it may have seemed to participants that the complexity of the task is lower with the system with the ECA than with the system without the ECA.

On the learnability of the process of extracting information there was an effect of the order of presentation (F (1, 24) = 8.075; p < .01) and an interaction between the type of ECA and type of content (F (1, 24) = 8.075; p < .01). The significant interaction between type of content and type of ECA was further analysed using simple main effect analysis. It showed that the variation of content across the order conditions significantly influenced how the participants perceived the difficulty of learning how to extract information from the system without the ECA (F (1, 24) = 9.084; p < .05) but not from the system with the ECA. Participants perceived the process of extracting information from the system without the ECA as more difficult to learn with the simple content than the technical content. The variation of content did not impact on how the participants perceived the learnability of the process when using the system with the ECA. Most likely, the modalities used by the system with the ECA made it easier for participants to learn how to extract information from the system.

On the learning of the information presented by the systems there was an effect of the order of presentation (F (1, 24) = 7.032; p < .05) and an interaction between the type of ECA and type of content (F (1, 24) = 7.032; p < .05). A simple main effect analysis on the interaction between the type of ECA and the type of content, showed that the variation of the content across the order conditions significantly affected the participants using the system with the ECA (F (1, 24) = 6.251; p < .05) but not the participants using the system without the ECA. Participants perceived the technical content as easier to learn than the simple content when it was presented by the system with the ECA. Therefore, we argue that the system with the ECA not only provides a more consistent method of presentation (see "Performance Measures"), but it also has the potential to enhance information retention of technical cultural content. As discussed later in the comments, participants felt that both contents were difficult to memorize, which explains the overall high ratings for this item with both systems. However, it seems that participants may have felt that the system with the ECA renders the interaction smoother thus making it easier for them to retain such information. This provides some evidence that supports my enhanced retention of complex information

hypothesis ("*The enhanced retention of complex information*"). However as the objective measures failed to produce any significant results, it is hardly possible to draw any conclusions solely based on subjective evidence.

Finally, on the self-organization of the information task there was an effect of the order of presentation ($F (1, 24) = 5.481$; $p < .05$) and an interaction between the type of ECA and type of content ($F (1, 24) = 7.032$; $p < .05$). Participants perceived the self-organization requirements of the technical content as higher than the simple content, when each type of content was presented by the system with the ECA. On a contrary, they perceived the self-organization requirements of the technical content as lower than the simple content, when each type of content was presented by the system without the ECA. However, the simple main effects analysis failed to reach conventional significance levels for either the system with the ECA or the system without the ECA. A possible explanation is that the text used by the system without the ECA in the technical presentations was more natural for participants to read than watching an ECA on the screen giving information acting almost, but not perfectly, like an actual human being.

Curiously, the participants' retention performances do not follow the findings reported above. One would expect that since the system without the ECA renders the technical task less difficult and with less self-organization requirements, it would translate to enhanced retention performances with the system without the ECA. Then, if participants thought that it is easier to learn the technical information with the system with the ECA than with the system without the ECA their motivation should have resulted to enhanced retention performances. However, as it can be seen from Table 1, participants' retention performances when experiencing the technical content were similar with both systems (mean ECA absent = 21.8 vs. mean ECA present = 21.1) and improved when experiencing the simple content (mean ECA present = 25.1 vs. mean ECA absent = 36.8) with the system without the ECA. A possible explanation is that regardless of how the participants perceived the technical task with the system with the ECA, the modalities it used impacted their ability to memorise the content equally, as the modalities used by the system without the ECA.

Participant Comments: Participants were asked to comment freely on each of the systems. To analyse the gathered data a custom-made approach was used. In particular, the group of participants was divided into a "Feedback" and a "Confirmation" group. Each group consisted of an equal number of participants randomly chosen from the experimental conditions. If a participant had not provided feedback, s/he was excluded from the groups. We looked in the feedback group for comments based on:

- Frequency: These were groups of comments that frequently arose around a specific event encountered or feature of the systems (e.g. ECA design, experiences with the multimodal content, etc.). As a frequency threshold for these patterns, we defined 40% of the total number of participants in the feedback group.
- Fundamentality: These were comments that although did not frequently arise were deemed be fundamentally important in terms of the possible effects of ECA on the user's experience of the prototypes.

Then, we looked in the confirmation group for comments that corroborated the patterns and/or comments of the feedback group. If a match was found in the confirmation group then, the pattern/comment of the feedback group was considered as corroborated.

If a match was not found the pattern/comment of the feedback group was considered as uncorroborated. As participants in both groups were taken at random from the experimental conditions, views were mixed and therefore uncorroborated patterns/comments could arise in the analysis. Below, we report only the corroborated patterns and comments grouped into relevant topics for simplicity.

ECA Design:

1) Certain features of the avatar can be improved. These include:

- Decrease the rate of the ECA's speech to make memorization easier
- Better body gestures
- More natural voice to avoid the speech discrepancies.

This corroborated pattern, suggests a number of improvements to the design of the ECA that, if implemented correctly, they could make memorization of the content easier. Participants did not have any comments about the photorealism of the ECA, which leads me to assume that it was acceptable, and focused only on the improvement of the ECA's behaviours and voice.

Multimodal Content Design:

1) The content is difficult to comprehend and memorize for most users. This could be because of the nature of the content, as it was suggested by a participant. A content of more historical value and without so many dates (as opposed to information about the construction of the churches), could be of more interest to the users.

This pattern was corroborated by a number of participants. It shows that a different type of content (i.e., of a more historical value without so many dates) would be of more interest to the users. It also provides a possible explanation why participants scored overall low in the retention tests using both systems. A more personalised content to the preferences of participants may reveal stronger differences between the two systems (ECA-present and ECA-absent).

Application Design;

1) **New Features**

- Zoom-in options to better see the artefacts for which the system is providing information.
- A pause button to pause the presentation on demand.

The above corroborated pattern and comment reveal features that should be added to the systems. Participants requested a pause button and a zoom option for the artefacts the system is providing information about. Although the ECA pointed to the artefacts, the resolution of the background images was low. The zoom feature will most likely make comprehension of the narrated content easier.

2) Improvements in the existing design

- A number of participants had problems photographing the QR-Code

The comment above require improvements in the process of photographing the QR-Codes in the locations. According to the observations made by the experimenter, participants experienced two types of problems with the QR-Codes: a) some of the participants had problems photographing a QR-Code, as the lab was too bright. This can be solved by adding a higher resolution camera to the device. b) When the ECA was present, the experimenter noticed that the "click" sound of the system's camera did not work consistently. This was to be expected, as the UMPC device had to process the graphics of the avatar, in addition to the video stream needed to photograph the QR-Codes. The latest generation of UMPC devices offers significantly more processing and graphics power, than the device used in this experiments.

5. Conclusion and Future Work

This study examined the effect of an ECA compared with a text/voice system in helping users extracting content with varying degree of difficulty in a simulated outdoor environment. In this study, participants using a multimodal ECA performed more consistently with the content, that those who used the system with text and voice output. However, qualitative evidence indicates that an ECA with text as an additional output modality can positively impact the perception of the technical content. This can potentially lead to enhanced retention of technical information.

Qualitative results revealed that the cognitive load of participants, with regards to the organization and implementation requirements of the information task, was affected by the complexity of the content, and the type of system that presented it. The system without the ECA lowered the difficulty and self-organizational requirements of the technical task, but placed more demands to the participants in terms of its learnability. The system with the ECA, on the other hand, lowered the learnability of the technical task, but had a detrimental effect on the perception of its difficulty and self-organization. Although this finding can be generalised only with caution, the careful use of text as an additional output modality when presenting technical content in an ECA-based mobile guide system can be beneficial for users rather than degrading. Then, the use of QR-Codes as a technique of content-tagging random locations in an outdoor attraction was received positively by the participants even under simulated conditions. Although some participants experienced issues with photographing accurately a QR-Code the first time, they all managed to complete a full tour with both systems with no prior training.

Overall, this study provided quantitative and qualitative evidence to support the "Persona Zero-Effect" [3] that an ECA has no negative or positive impact on the users experiencing cultural content in simulated mobile conditions. However, my results also suggest that by augmenting the output modalities of an ECA with text when presenting technical content can be beneficial for users rather than degrading.

6.   References

[1] Apple Inc., (2013). "Siri Personal Assistant". (Available from: http://www.apple.com /ios/siri/) [Accessed February 19 2013].

[2] Lim, Y., Aylett, R. (2007): "Feel the difference: a guide with attitude!" In Pelachaud, C., Martin, J.C., Andre, E., Chollet, G., Karpouzis, K., Pele, D., eds.: *Proceedings of the 7th International Conference on Intelligent Virtual Agents (IVA)*, Springer (2007) pp. 317 – 330.

[3] Miksatko, J., Kipp, K.H., Kipp, M. (2010): "The Persona Zero-Effect: Evaluating virtual character benefits on a learning task." In: *Proceedings of the 10th International Conference on Intelligent Virtual Agents (IVA-2010),* September 20-22, 2010, Philadelphia, PA, USA, pp. 475-482.

[4] MacDorman, K. F., Gadde, P., Ho, C.-C., Mitchell, W. J., Patel, H., Schermerhorn, P. W., & Scheutz, M. (2010): "Probing people's attitudes and behaviours using humanlike agents." *Research Poster IUPUI Research Day*. April 9, 2010. Indianapolis, Indiana.

[5] Hollan, J., Hutchins, E., & Kirsh, D. (2000): *"Distributed Cognition: Toward a New Foundation for Human Computer Interaction Research".* ACM Trans. on *Computer-Human Interaction*. **7 (2)**, pp. 174-196.

[6] Nemetz, F., & Johnson, P. (1998): "Towards Principled Multimedia" [online]. University of London. Available from: http://www.bath.ac.uk/~mapfn/chi98/ towards.htm [Accessed July 15 2011].

[7] Baylor, A. L. and S. Kim (2008): "The Effects of Agent Nonverbal Communication on Procedural and Attitudinal Learning Outcomes". *International Conference on Intelligent Virtual Agents*. Tokyo, Springer**:** pp. 208-214.

[8] Likert, R. (1932): "A Technique for the Measurement of Attitudes". In: *Archives of Psychology* **140,** p.55.