Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 423-430

Attentional Effect of Animated Character

***Cholyeun Hongpaisanwiwat & **Michael Lewis**

*Department of Computer Science, Thammasat University,

99 Mhu 18 Paholyothin Rd, Khong Luang, Prathum Thani 12121, Thailand

**School of Information Sciences, University of Pittsburgh,

135 N. Bellefiled Ave, Pittsburgh, PA 15260, USA

*chhong@cs.tu.ac.th, **ml@sis.pitt.edu

Abstract: Research has found that animated characters are capable of capturing users' attention, engaging them in active tasks, and entertaining them. Such capabilities are pedagogical techniques that might contribute to an effective comprehensible multimedia presentation. Quality of voice has also been shown to be an important determinant of comprehension. The current study examines the effects of animated characters and voice types on comprehension and attention performance in learning from a multimedia presentation. Because animated characters rely on affective social responses to produce pedagogical benefits, there are also likely to be significant individual differences. This study investigates the influence of introversion/extroversion on the effectiveness of the presentations as well. A 3 (no character, nonanthropomorphic character, anthropomorphic character) x 2 (synthetic voice vs. human voice) factorial between-subject design was employed. While animated characters did not increase learning in the present study; results suggested that animated characters might be useful as a peripheral tool to retain learners' attention and maintain engagement with the learning material. Results regarding the degree of participants' extroversion (benefits appeared limited to introverted participants) suggest that personality should be taken into account in designing learning environments.

Keywords: Animated character, comprehension, attention, multimedia presentation, extroverts and introverts

1 Introduction

Although the value of animated characters in user interfaces has not been established, many believe such characters might be useful in educational software. If animated, anthropomorphic agents can mediate communication between computers and users, then human-computer interaction becomes more similar to human-human communication. These effects could be useful in learning environments. Therefore, animated agents are often proposed as a way to improve students' learning performance. In this paper, the motivation which leads to hypotheses is first presented, following by research method, research results, and conclusion and discussion.

2 Motivation

Drawing attention and active engagement are pedagogical techniques for designing effective multimedia presentation. Generally, both visual and audio cognitive resources are occupied while learning from a multimedia presentation. When presentations are complex, these channels can become overwhelmed and it may be necessary to guide attention to important elements. This is particularly important for younger learners or persons with low domain knowledge who may have difficulty in discriminating the relevant and irrelevant materials (Miller & Weiss, 1981). The method for directing attention may be as simple as a pointer used to reference significant elements in the display. Faraday and Sutcliffe (1997) found the use of an arrow effective in controlling or guiding attention within a

presentation; All of the subjects in their study shifted their attention as an arrow pointed to various objects. Additionally, they found that the objects to which the subjects' attention was directed were more accurately recalled.

The value of directing attention is based on cognitive information processing theories which hold that attended information consumes more cognitive resources and is more thoroughly processed than unattended information. Researchers report that anthropomorphic computer interfaces have more capabilities of drawing users' attention and engaging them in active tasks than nonanthropomorphic interfaces (Dehn, & van Mulken, 2000). Therefore, human-like characters could be more effective presenters or guiders in multimedia presentations where comprehension and recall of materials need to be enhanced. Therefore, the users viewing a presentation with an animated character, having expended greater resources and processing effort on the course material should better understand the material than the users who view the presentation without such a character.

Nonetheless, users while interacting with a human face lost concentration on their task because of attempting to interpret the facial display (Takeuchi & Naito, 1995). A similar effect could occur in educational multimedia especially, in a learning environment of complex multimedia presentations where cognitive resources are required to process audio and visual information. If the animated characters become a source of distraction, they could impede a student's comprehension of a presentaion rather than enhancing it. Clearly, this issue needs further investigation before animated agents are accepted by educational multimedia software packages.

In addition, quality of voice is an important determinant of several factors including comprehension (Lai, Wood, & Considine, 2000; Sonntag, Portele, & Haas, 1998; Rosson & Cecala, 1986), of a user's perception of characters (McBreen & Jack, 2000), of characters' intelligibility (Beskow, 1997), and of emotion and personalities (Ekman, Friesen, O'Sullivan, & Sherer, 1980). Although advanced synthetic speech, known as text-to-speech, has been used in many applications such as telephony systems and reading electronic mail, a human voice is generally preferred. Research has found that while comprehension of synthetic speech was significantly worse than human recorded speech, after people became used to the synthetic speech, the comprehension performance of both speech conditions (human-recorded and synthetic) appeared comparable (Lai, Wood, & Considine, 2000). In addition, with respect to the consistency theory, the mismatch condition of a talking head with human voice was less socially appealing than the match of a talking head with synthetic voice (Nass & Gong, 2000). This indicates that an animated character with a human voice may be paid less attention to than an animated character with a synthetic voice. Therefore, a presenting voice for a human-like presenter should impact learners' comprehension of and attention to presented materials. The current study explores the effect of the manipulation of the voices of animated characters.

The focus of this study is to examine the effects of anthropomorphic animated character and voice on comprehension and attention in multimedia presentations of technical content. Human forms compared to others representations (such as symbolic shapes and Chernoff faces) were judged as the greatest degree of anthropomorphism (King & Ohya, 1996).

Furthermore, emotions and social behaviors are necessary for life-like perception (Laurel, 1997; Maes, 1995). Therefore, in this study, an *anthropomorphic character* was defined as a human-like representation, which is capable of displaying emotional expressions, while a *nonanthropomorphic character* was an object or a part of human form, which was not able to express emotions.

The study measured the degree of attentiveness attributed to animated characters by their effects on incidental learning, measured as the accuracy of recall performance for attended visual messages and the number of errors in recall performance for unattended visual messages. If an anthropomorphic character elicits and holds user's attention to a greater degree than a nonanthropomorphic character then users should have better recall for attended messages and poorer recall for messages presented elsewhere on the display. With respect to the quality of the presenting voice, a human voice is expected to produce greater learning comprehension and attention performance than a synthetic voice. This leads to the following hypothesis.

Hypothesis 1: Students who learn a multimedia lesson with different conditions of animated characters and voice types will have different comprehension and attention performance. The interaction effect of anthropomorphic character and voice types could be expected due to the effect levels of these two factors. It was expected that differences in comprehension mean and attention mean between synthetic and human voices in the

anthropomorphic animated character condition would be the greatest, followed by differences in the nonanthropomorphic animated character condition, and that the differences in the condition without animated characters would be of least significance.

Additionally, individual differences are known to be a significant predictor in human computer interaction. This study also examined the effect of animated characters as a function of the participant's degree of extroversion. The extroversion/ introversion dimension is closely related to social interaction. It could be the source of pedagogical agent effects; thus, extroversion scores were included as an additional factor. According to research findings in educational psychology, extroverts and introverts have different social interaction styles and prefer different learning environments (Fairhurst & Fairhurst, 1995; Jonassen & Grabrowski, 1993). It can be hypothesized following:

Hypothesis 2: Extroverts and introverts will have different performances on both comprehension and attention.

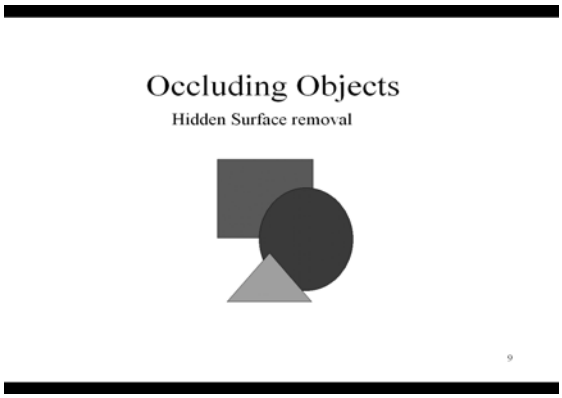
3 Methods

3.1 Manipulation: Animated Character

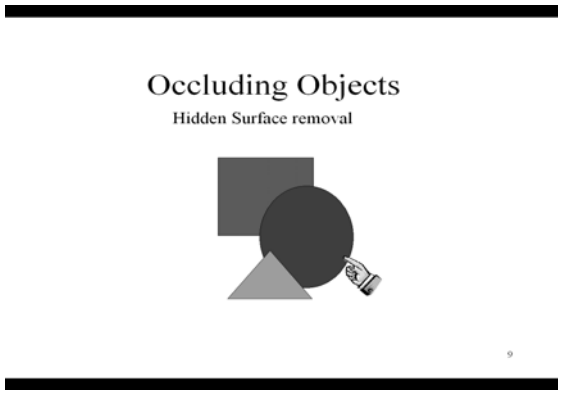
In an effort to simulate a real-life situation, in this study the PowerPoint slides used in an actual computer graphics course were used with the different attention directing presentations. The lecture introducing the concept of three-dimensional computer graphics was recorded and transcribed. The presentation consisted of 33 slides. The transcription was divided by slide. The script for each slide comprised of both the instructor's explanation and a sequence of locations to which attention should be directed. The 33 slides were divided into two sets. The first set covered the overview of the five notions to create three-dimensional illusion and the second set presented the first two notions (perspective and occlusion) in greater depth, including computational detail.

Figure 1: Animated character: Finger and Paul

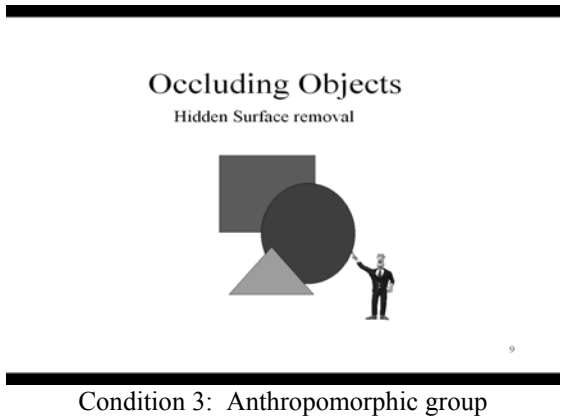
Two animated characters were used in this study (See Figure 1). A finger was used to represent a nonanthropomorphic character, and the character named Paul represented an anthropomorphic character. Both characters had the same scripts for pointing behaviors (for an example, see Figure 2 screen shots), but the anthropomorphic character could express other behaviors such as denying, giving suggestions (See examples in Figure 3).



Condition 1: Control group – no character



Condition 2: Nonanthropomorphic group



Condition 3: Anthropomorphic group

Figure 2: Screenshot examples of pointing behaviors

Figure 3: Examples of Paul's behaviors

The manipulation of the characters accounted for three conditions in this study: 1) multimedia presentation presented without an animated character--control group, 2) multimedia presentation with the finger--nonanthropomorphic group, and 3) multimedia presentation with Paul--anthropomorphic group.

3.2 Manipulation: Presenting Voice

Both a real human voice and a synthetic voice were used in the experiment. A human voice from the real lecture was edited in order to eliminate non-speech expressions such as "er…..uh" as well as background noises so that speech time remained close to text-to-speech.

3.3 Measure

The dependent variables including comprehension and attention were measured using a set of paper-and-pencil tests. The accuracy of these tests was used to assess comprehension. The measure of attention was determined by the number of accurate items of recall on objects to which attention was explicitly directed (Pointed-to) and the number of errors of recall items of objects to which attention was not explicitly directed (Not Pointed-to).

3.4 Experimental Design

In this study, there were two independent variables: 2 levels of narration voices and 3 levels of animated characters. The no presenting character level was the control condition. The experiment design was a 2 x 3 between-subject design illustrated in Table 1.

Presenting Voice	No Character (Control)	Non Anthropo-morphic character (Finger)	Anthropo-morphic character (Paul)
Synthetic	Group 1	Group 2	Group 3
Human	Group 4	Group 5	Group 6

Table 1: 2 x 3 Between-subject Design

3.5 Participants

Sixty students at the University of Pittsburgh who had not taken computer graphics were recruited as participants. All participants were Native English speakers. They were randomly assigned to one of the six experimental conditions and were paid ($15) for participation in the study. Each condition was comprised of ten participants.

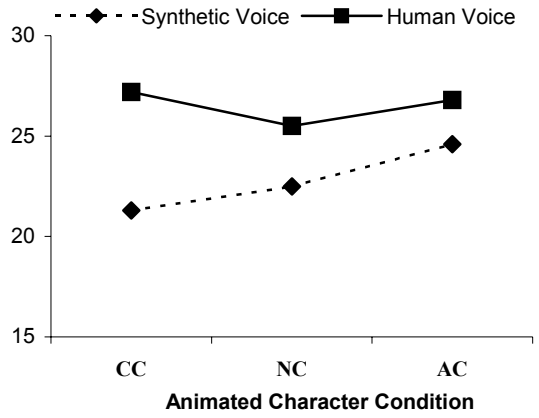
Of the 60 participants, 48% were male and 52% were female. Most of the participants were undergraduate students who came from a variety of disciplines, primarily the area of science. The ages of participants ranged from 18 to 46, with a mean of 21.4. The participants' extroversion scale ranged from 19 to 70, with a mean 51.25 and a standard deviation of 9.38.

3.6 Procedure

Participants were randomly assigned to one of the conditions. Upon arrival, they were asked to complete the online Extroversion Inventory developed by Plumeus Inc. (1996). Then they were told to view two sets of slide presentations. The presentations were set to run automatically without any intervention from the students. The participants were required to wear the headphones and participants who were assigned in synthetic voice conditions were informed that they would be listening to a computer-generated voice for the presentation. When the first presentation stopped, the participants were given a test before proceeding to the second presentation and finally were given other two tests.

4 Results

4.1 Effects on Comprehension and Attention Performance

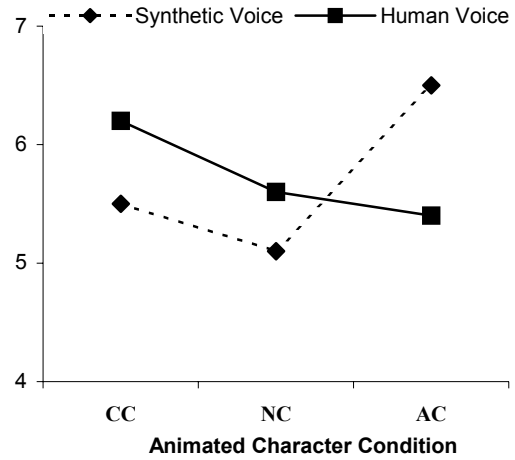


CC = Control Condition,
NC = Nonanthropomorphic Condition
AC = Anthropomorphic Condition

Figure 4: Comprehension performance

The comprehension performance was analyzed by analysis of variance with presentation treatment (animated character condition x voice type). The results indicated that there were no significant main effects of animated character conditions and voice types, and no significant interaction effects between animated character conditions and voice types regarding comprehension. However, voice type did marginally affect the comprehension performance [$F(1,54)=3.344$, $p < 0.073$]. As shown in Figure 4, the human voice condition produced marginally better comprehension than did the synthetic voice.

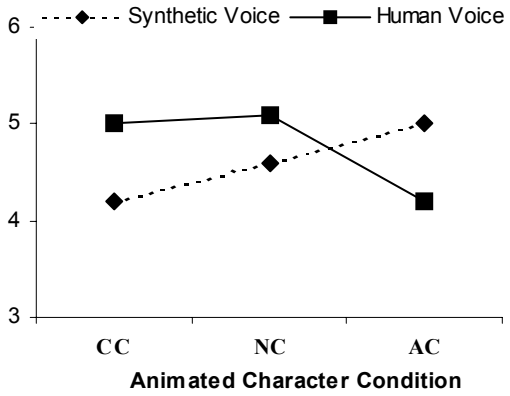
Attention was measured within this study by contrasting recall performance for pointed-to (attended) and non-pointed-to (unattended) objects.



CC = Control Condition,
NC = Nonanthropomorphic Condition
AC = Anthropomorphic Condition

Figure 5: Recall performance for pointed-to (attended) objects

The results of an analysis of variance indicated that neither animated characters nor voice types had a significant effect on recall performance for "pointed-to" objects; however, the interaction of the two did have effect on recall performance. This indicated that the animated characters and voice types influenced subjects in recalling "pointed-to" objects differently. As shown in Figure 5, the anthropomorphic animated character in the synthetic voice condition was helpful in recalling pointed-to objects, but for the human voice condition, better recall performance was not increased. In addition, this figure indicates worse recall performance for anthropomorphic animated character in the human voice condition.



CC = Control Condition,
NC = Nonanthropomorphic Condition
AC = Anthropomorphic Condition

Figure 6: Recall performance for non-pointed-to (unattended) objects

Similar to the results of recall performance for "pointed-to" objects, there were no significant main effects of the animated character condition and voice type on recalling "non-pointed-to" objects, and there was no interaction effect as well. Figure 6 shows that using (non)anthropomorphic animated characters did not cause a poorer performance in recalling the objects which were not explicitly pointed to in the presentation. Participants who experienced the (non)anthropomorphic animated characters with any voice type could recall the "non-pointed-to" objects with an accuracy equal to that of the participants in the control groups. Regardless of animated character condition, participants who experienced the human voice did not perform better in recalling the "non-pointed-to" objects than those participants who heard the synthetic voice.

The first hypothesis was accepted because results from the experiment indicated that the animated characters and voice types did not affect the learners' comprehension. However, attentional performance in recalling "pointed-to" objects was affected by the animated characters and voice types that were used. Attention for "non-pointed-to" objects was not affected either by the animated character or voice type.

4.2 Influence on Extroverts and Introverts

There are fewer introverts than extroverts in general. Participants were divided into unequal groups of low and high extroversion, by using 50 (the median from the Extroversion Inventory's validation sample) as the cutoff. Twenty-five participants (42%) were

classified as introverted while thirty-five (58%) were classified as extroverted.

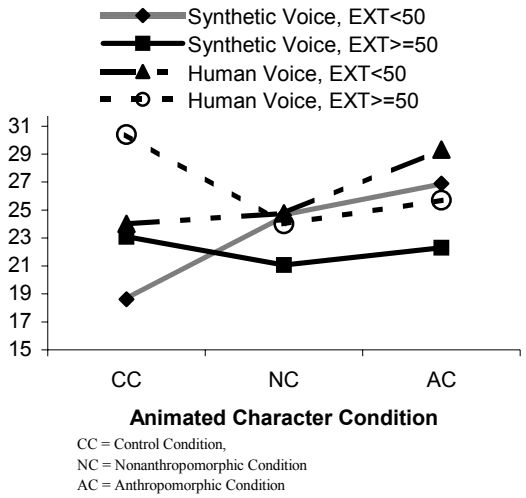


Figure 7: Comprehension performance by extroversion scale

The results of a three-way analysis of variance indicated that there were no significant main effects of animated character condition, voice type, or extroversion scale. However, the effect of voice type approached significance [$F(1,48)=3.80$, $p = 0.057$]. As illustrated in Figure 7, those participants who viewed the presentation with a human voice comprehended the material better than those who viewed the presentation with a synthetic voice. There were neither two-way nor three-way interaction effects to report.

Interestingly, although insignificant, the two-way interaction between animated character conditions and extroversion scale demonstrated that participants who were categorized as less extroverted tended to comprehend the material better with animated characters, while those participants who leaned toward extroversion tended to learn better without animated characters.

The results of a three-way analysis of variance of recall involving "pointed-to" objects indicate that there were no main effects of animated character condition, voice type, or extroversion; however, there were two two-way interactions and three-way interaction effect. The animated character condition by voice type interaction was significant [$F(1,48)=6.531$, $p = 0.003$] and the other two-way interaction effect was between the voice type and the extroversion scale [$F(1,48)=6.405$, $p = 0.015$], but the animated character condition by the extroversion scale interaction was not significant.

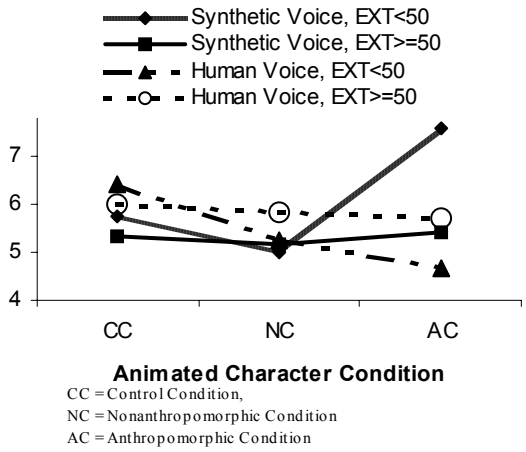


Figure 8: Recall performance for pointed-to (attended) objects by extroversion scale

Figure 8 shows the effects of interaction for recall involving "pointed-to" objects between human and synthetic voices. The greatest differentiation took place in the anthropomorphic animated character condition, followed by the condition without an animated character. The least differentiation took place in the nonanthropomorphic animated character condition. While the performance for the nonanthropomorphic animated and no-animated character with human voice was higher than the performance which featured the synthetic voice, performance in the condition which featured the anthropomorphic animated character with synthetic voice showed much higher than results the condition which featured the human voice.

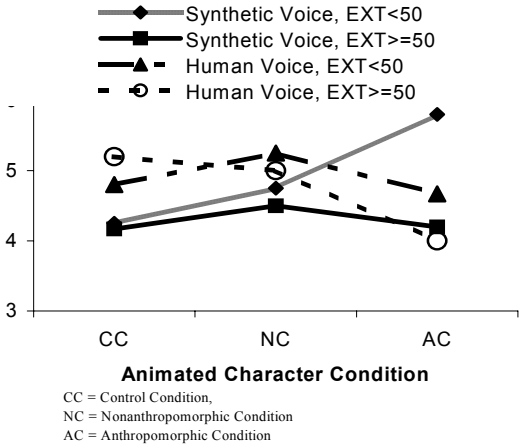


Figure 9: Recall performance for non-pointed-to (unattended) objects by extroversion scale

The recall performance involving "non-pointed-to" objects was analyzed using analysis of variance with treatment (animated character condition x voice type) by the extroversion category. There were no significant main effects of animated character condition, voice type, or extroversion scale to report, nor were there any significant two-way or three-way interaction effects.

Hypothesis 2 predicted that extroverts and introverts would have different comprehension and attention performances. The results show that there were no apparent effects of animated characters and voice types in regard to extroversion scale when employed as an additional between-subject factor on comprehension and attention performance or on rating the presentation and animated characters. Both extroverted and introverted learners comprehended presentation material equally for all treatment conditions. Attention performance for recall involving "pointed-to" objects depended upon animated character condition, voice type, and degree of extroversion; however, for extroverted and introverted learners in all treatment conditions, attention performance with respect to recall involving "non-pointed-to" objects was equal.

5 Discussion and Conclusion

The current study investigated the use of an animated character in the role of a presenter of a multimedia presentation. Its objective was to determine to what extent animated characters could hold students' attention during a presentation so that students' comprehension could be enhanced. The study found that the presence of an animated character as a presenter had no impact on the comprehension of the presentation. These findings are consistent with the work by Andre et al. (1998) in which the presence of an animated agent made no difference to participants' comprehension of a Web-based presentation.

There was concern that the animated character might prove to be a distraction for the participants. This study found that the anthropomorphic animated character was marginally less distracting, although more engaging than the simple animated character. These findings also correspond to reports (Andre, Rist, & Mueller, 1998), which found that an animated agent was not distracting.

The present study provided evidence supporting the findings (Sproull et al, 1996; Takeuchi & Naito, 1995) that an anthropomorphic computer-interface that includes human facial displays or synthetic animated characters can draw users' attention to current tasks. However, the consistency of an animated character with its voice must be taken into account in order to draw student's attention; a match between a synthetic character and a synthetic voice produced greater recall performance than a mismatch between a synthetic character and a human voice. The consistency effect found in the current study supports earlier findings that a mismatch of a synthetic talking head with a natural voice has less social appeal than that of the match of a synthetic talking head with a synthetic voice (Nass & Gong, 2000).

The results of this study failed to provide evidence that using anthropomorphic animated characters as presenters in a multimedia-learning environment has any impact on student learning. This result corresponds to the previous studies (Andre, Rist, & Mueller, 1998; Okonkwo & Vassileva, 2000) in which the personified characters did not contribute to the students' comprehension of technical materials. However, the anthropomorphic character was better able to draw student's attention to the messages that were being presented than the nonanthropomorphic character. Although attention is critical in learning, it is not always related to knowledge transfer. In further analysis, no correlation between attention and comprehension was found. Therefore, anthropomorphic animated characters do not appear to help learners digest presented materials, but may be an effective peripheral learning tool for directing attention where needed.

This study shows that the human voice did not produce better comprehension than the synthetic voice. This finding may indicate that a synthetic voice has been improved and its quality becomes more comparable with a human voice. Therefore, a synthetic voice may be used to author a multimedia presentation package without loss of usability. This is a good news for multimedia lessons because authoring a multimedia learning package with a human voice narration takes more time, is more expensive, and more difficult to maintain.

The current study found a significant aptitude-treatment interaction (Cronbach & Snow, 1977) between individual differences of learners and learning environments. Participants in the low extroversion group tended to have greater comprehension when viewing the multimedia presentation narrated by an animated character, while participants in the high extroversion group tended to have better comprehension scores when viewing the multimedia presentation without an animated character narration. This may have occurred because extroverted learners may not

attend well in non-interactive environments since they prefer non-structured materials with lots of collaborative activities and interactivity; introverts, on the other hand, benefited from the structured material and quiet learning environment (Jonassen & Grabowski, 1993).

This study suggests that designers of instructional multimedia cannot safely disregard individual differences among their audiences. Our results indicate that customizable learning environments that allow the student to selected preferred instructional modes or adapt to match student characteristics may hold the greatest promise. The significant three-way interaction among animated characters, voice types, and degree of extroversion for recall "pointed-to" objects indicates that learner extroversion is a salient variable which should be more widely considered in educational multimedia research.

References

Andre, E, Rist, T, & Mueller, J. (1998), WebPersona: a lifelike presentation agent for the World-Wide Web, *Knowledge Based Systems*, 2 (1), 25-35.

Beskow, J. (1997), Animation of talking agents. *In Proceedings of AVSP'97, ESCA Workshop on Audio-Visual Speech Proceeding*, Rhodes, Greece.

Cronbach, L. J., & Snow, R. E. (1977), *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. New York, NY: Irvington Publishers, Inc.

Dehn, D. M., & van Mulken, S. (2000), The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies,* 52, 1-22.

Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. (1980), Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology*, 38(2), 270-277.

Faraday, P. & Sutcliffe, A. (1997), Designing effective multimedia presentations. *In Proceedings of CHI'97 Conference Compassion,* ACM Press, 272 - 278.

Jonassen, D. H., & Grabowski, B. L. (1993), *Handbook of individual differences, learning & instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

King, W. J., & Ohya, J. (1996), The Representation of Agents: Anthropomorphism, Agency, and Intelligence. *In Proceedings of CHI'96*, ACM Press, 289-290.

Lai, J., Wood, D. & Considine, M. (2000), The effect of task conditions on the comprehensibility of synthetic speech. *In Proeedings of. CHI 2000*, ACM Press, 1-6.

Laurel, B. (1997), Interface agents: Metaphor with character. In J. M. Bradshaw (Ed.), *Software Agents*, Cambridge, Massachusetts: MIT Press, 67-77.

Maes, P. (1995), Artificial life meets entertainments: Life-like autonomous agents. *Communication of the ACM*, 38(11), 180-114.

McBreen, H. & Jack, M. (2000), Animated Conversational Agents in E-Commerce Applications. *Proc. Third Workshop on Human Computer Conversation*, pp. 112-117. Bellagio, Lago di Como, Italy.

Miller, P. H., & Weiss, M. G. (1981), Children's attention allocation, understanding of attention, and performance on the incidental learning task. *Child Development*, 52, 1183-1190.

Nass, C. & Gong, L. (2000), Social aspects of speech interfaces from an evolutionary perspective: Experimental research and design implications. *Communications of the ACM*, *43*(9), 36-43.

Okonkwo, C., Vassileva, J. (2001), Affective Pedagogical Agents and User Persuasion in C. Stephanidis (ed.) *Proc. "Universal Access in Human - Computer Interaction (UAHCI)"*, held jointly with the 9th International Conference on Human-Computer Interaction, New Orleans, USA, 5-10 August 2001.

Plumeus, (1996). http:://www.queendom.com.

Rosson, M., & Cecala, A. (1986), Designing a quality voice: An analysis of listeners' reactions to synthetic voices. *In Proceedings.of CHI'86*, ACM Press, 192-197.

Sonntag, G. Portele, T., & Haas, F. (1998), Comparing the comprehensibility of different synthetic voices in a dual task experiment. *Proceedings of the Third ESCA/COCOSDA Workshop on Speech Synthetic*, Blue Mountain, Australia.

Sproull, L., Subramani, M., Kiesler, S., Walker, J.H., & Waters, K. (1996), When the Interface is a Face. *Human-Computer Interaction*, 11, 97-124.

Takeuchi, A., & Naito, T. (1995), Situated facial display: Towards social Interaction. *In Proceedings of CHI'95*, ACM Press, 450-455.