# User Satisfaction of Non-Realistic Three-Dimensional Talking-Head Animation Courseware (3D-NR)

Mohd Najib Hamdan*, Ahmad Zamzuri Mohamad Ali

Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, Malaysia.

**Abstract:** Talking-head animation is an instructional approach that helps students in linguistic learning, especially in the pronunciation aspect. However, the use of talking-head animation has caused some emotional uneasiness among students when the design and realistic level of the animated characters are too human-like. This phenomenon is known as 'Uncanny Valley' and was corroborated through a research by a Japanese robotics expert named Masahiro Mori in 1970, who had also produced a graph explaining this phenomenon. To overcome these problems, the non-realistic three-dimensional talking-head animation has been developed to ensure that students get the maximum learning from the point of emotional and learning performance. On this regard, the focus of this study was mainly on developing non-realistic three-dimensional talking-head animation. The courseware was developed based on theories, principles and literature overview conducted. The paper also reports the outcome of usability and user satisfaction (PSSUQ) test carried out.

**Key words:** Animation, emotion, instructional, realism, talking-head.

## 1. Introduction

Multimedia, especially animation applications are frequently used as a Teaching Aid in the modern education [1], as the elements found in the animation make teaching and learning more attractive, effective, stimulating, and it can be suited according to the needs of students [2]. Currently, animation is not only used in certain fields or discipline such as science and technology, but it has also been expanded into the fields of language and linguistics [3]. This includes animation in the form of talking-head which is being used in learning language and linguistics [4], [5]. It is found that by using the talking-head animation, students can learn a language by imitating the style and speech according to the animated character [5]. This argument has been proven by several researches related to learning to pronounce words in English where there is a positive effect on students who used the talking-head animation [5], [6] Therefore, the use of talking-head animation has the potential to become a Teaching Aid in the learning of words, which is the main problem in learning English as a second language [7]. However, the effectiveness of the talking-head animation is closely related to the character used in the animation [8].

Generally, this animated character can be divided into two categories, which are the two-dimensional (2D) and three-dimensional (3D) characters [9]. The three-dimensional character, which is very human-like, is said to have an emotional effect on students during the learning process compared to the two-dimensional animated characters [10], [11]. This is due to the realism factor found in the character itself [12]. Thus, if this realistic factor is not handled properly, it could have a negative impact on the learning process [12].

## 2. The Effect of Realistic Levels of Animation Characters and Uncanny Valley Phenomenon on Human

The talking-head animation has a potential to involve the use of visual and verbal channels during the learning process, whereby the facial expressions and lip movements in the talking-head animation will be processed through the visual channel, whereas the texts and audio are processed through the verbal channel. Hence, effective learning may be interrupted if an unsuitable animation character design is used, although student's cognitive burden can be reduced by using both visual and verbal channels. This is because the realistic factor of an animation character will have an effect on student's emotions and level of comfort during the learning process, especially by using extremely realistic characters [12]. This effect may also have an impact on the success of instructional materials based on animation that is used in learning. In relation to that, it is important that research is carried out to determine the effect of the level of realism in animation characters on learning.

The effect that is created from the realistic levels of an animation character can be seen in the Uncanny Valley phenomenon. The Uncanny Valley phenomenon explains the relationship between the differences in the realistic level of a character, as well as human comfort and emotions [13]. This theory elucidates the relationship through the Uncanny Valley graph as shown in Fig. 1.
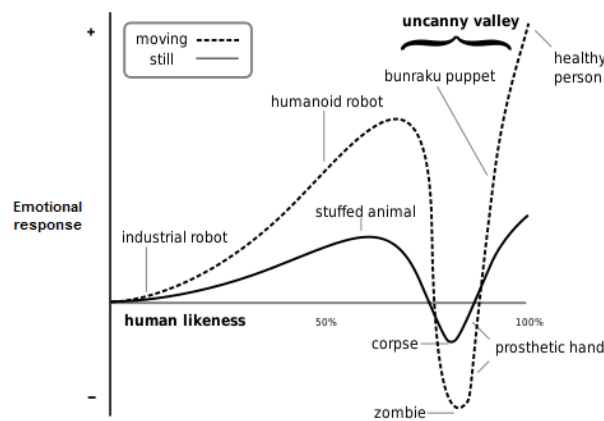


Fig. 1. The uncanny valley phenomenon [14].

Based on Fig. 1, the Uncanny Valley phenomenon divides the character into two main categories; dynamic characters and static characters. The findings of the research that have been converted into a graph show that human emotions increased positively when the character resembles a human. The graph shows that the level of human emotions decreased to a negative level until 80 to 85 percent for both the dynamic and static characters as the characters become more horrifying. The decrease in the level of human comfort in the graph is called the Uncanny Valley. Uncanny Valley also imparts a horrifying and frightening effect to viewers according to the design and expression of the character's face that is increasingly realistic and resembling humans [10]. The graph on the level of human emotions increases to a positive level when the character is an actual human [14]. Based on this graph, it was also found that the Uncanny Valley effect was higher with dynamic characters compared to static characters.

In view of this, the design and development of the talking-head animation for this research used non-realistic three-dimensional animated characters to avoid the Uncanny Valley phenomenon befalling the students. The findings of this study could help to determine whether the design of this research is similar to the standard of an animated instruction, specifically for learning to pronounce English words.

## 3. The Design and Development of the Talking-Head Animation

The process of creating and developing an effective multimedia learning material needs an instructional design model that is well arranged to ensure a systematic development process [15]. This matter is no exception in the process of developing instruction materials based on the talking-head animation, which also needs a

pertinent instruction design model in order to have a maximum impact in terms of enhancing the performance of students. According to "[16]", the evaluation and implementation involve almost 40 instruction material design models. Hence, most of these models consist of the same elements such as analysis, design, evaluation and implementation [17]. Among the instruction material design models are the ADDIE model, the Assure Model, the Product Orientation Model and the Dick & Carrey Teaching Design Model [18]. Thus, it is found that most of the novice developers of instruction materials faced problems in establishing the assignments in each phase of the instruction design model because the framework that is presented is only represented by words [15]. To overcome this problem, the DIDEA instruction design model was developed by Universiti Pendidikan Sultan Idris (UPSI) with the aim of being a guide on the procedures of developing the instruction material involved in the teaching and learning process [15]. DIDEA is the acronym for five phases, which are Determine, Illustrate, Development, Execute and Analyze. Each phase in the DIDEA model contains several sub-assignments that would assist in elucidating the function of each phase, and the sub-assignments are not present in most of the current instruction design models. Therefore, the DIDEA instruction design model has been selected in the designing and developing process of the talking-head animation in this research.
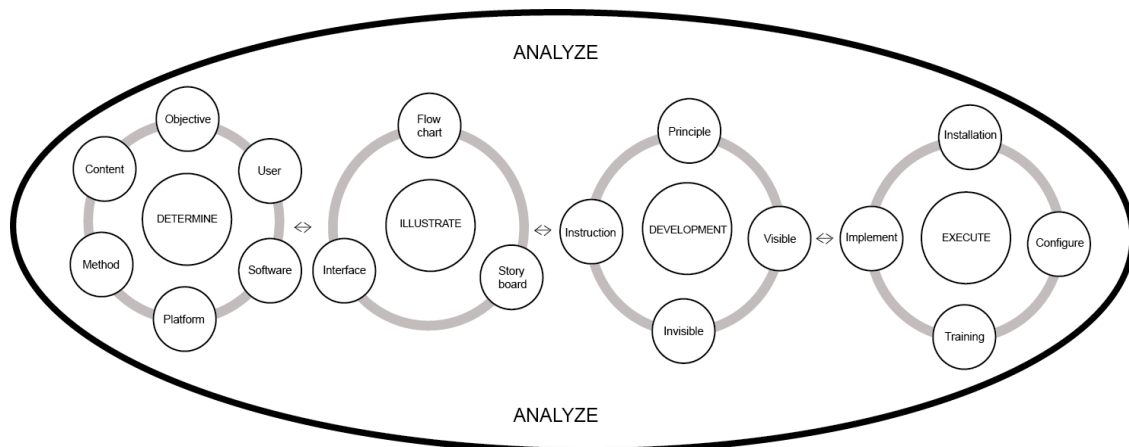


Fig. 2. The DIDEA model [15].

## 4. Navigation

The talking-head animation courseware begins with the introduction screen. From the introduction screen, students can choose to move directly to the main menu screen, the guide menu or the courseware objective menu. Through the main menu screen, they can choose a word to begin their pronunciation training. After the students have chosen a word, a screen will display the talking-head animated character that pronounces the same word with a display of the word in syllable break form (Example - Pro/nun/ci/a/tion). After that, the students will again look at the talking-head animated character pronouncing the word in full (Example - Pronunciation). Then, the students will look at the explanation on the meaning of the word shown on the next screen. After the learning session, they have the option to repeat the learning session for the same word or choose to learn another word.

Besides that, the talking-head animation courseware that has been developed involved the combination of linear and hierarchal navigation. The linear navigation is a structure that has a chain of continuous stories [19]. This enables students to follow the steps one-by-one from the beginning when they choose the word until the animated character pronounces the word. Meanwhile, the hierarchal navigation lends an opportunity for students to return to the main menu in order to select a new word or repeat the word that has been taught [20]. This supports the drill learning strategy method. The efficacy of the talking-head animation courseware is closely related to the principles of screen design [21]. According to "[15]", the development phase, especially the screen design aspect, comprises several multimedia principles that need to be fulfilled, such as principles

regarding texts, graphics, audio, animation, video, and colour.

In relation to that, the above mentioned multimedia principles should follow the knowledge discipline stated in the Human-Computer Interaction (HCI) field. This is intended to ensure the existence of a relationship or effective interaction between the consumer and the instruction material that has been developed in order to ascertain the effective delivery of information to the consumer [22]. The following are multimedia principles that have been practiced in the talking-head animation development phase.

## 5. Multimedia Principles

### 5.1. Text

Text is an arrangement of alphabets that form a clear expression and later leads to a meaning [21]. It comprises several combinations of symbols such as alphabets, numbers, types of fonts, notations, and styles of fonts [22]. Texts that portray facts or ideas of an individual are built from a combination of sentences, words and paragraphs [21]. Hence, text selection is important to ensure all the information is effectively delivered to consumer [21]. Generally, text is categorized into three main groups, which are the Serif type (Times New Roman, Bookman Old Style & Courier New), the San Serif type (Arial & Verdana) and the Decorative type (Old English Text & Egyptienne). The Serif type is text that is "tailed" at the end. Meanwhile, the San Serif that has endings that are not decorated [22]. The last group of texts is the decorative type, whereby the style of the text is in the decorative or flowery form. Selecting and using the decorative font would make it difficult to read, thus, a font that is easy to read and with an appropriate size is encouraged when developing a multimedia application [23]. Therefore, a textual concept that is brief and concise needs to be assimilated into the multimedia design [22]. The justification for a concise and brief text is that it is appropriately less than half of the screen size used [21]. Hence, when developing the talking-head animation courseware, the Arial font (San Serif) with an appropriate size was used for the entire software based on the advantage of using this font, such as ease of reading on the computer screen [22]. In this courseware, the texts are presented in a combination of large and small alphabets. Using large alphabets would make it difficult and impede the reading fluency because the words are all of the same size [24]. Fig. 3 shows an example fonts used in the courseware talking-head animation.



Fig. 3. Example of using the arial font and graphics in the talking-head animation courseware.

### 5.2. Graphics

Graphics are the main element in multimedia courseware that is used to process the visual delivery of information. Compared to the conventional method, information that is received through the graphical process is faster and easily received by students [25]. Good quality and effective graphics are capable of enhancing motivation and stimulates students to further focus on the learning sessions. Hence, using graphics that are too

decorative have the potential to disrupt student's concentration towards the learning material [26]. Therefore, graphic components are only used as icons, such as the "main page" (home), "next", "previous", "play" and "exit" so that it is uniform and easily understood by all students.

## 5.3. Audio

Audio is among the important elements used in this research to ensure that the words pronounced by the character are accurate and clearly heard by the user. Hence, every word pronounced by the talking-head character is based on the voice of a male character called Mike (a native speaker), which was obtained directly from a software known as 'text to speech' found in the website http://www.yakitome.com. Therefore, the voice of a local English teacher was chosen to pronounce the selected words, and this was verified by three lecturers who are experts in English. The audio was saved in the form of file.mp3 before it was used in the talking-head animation

## 5.4. Animation

Choosing the animation character is the most important component in the talking-head animation courseware. Hence, the choice of the character was evaluated based on typical Malaysian characteristics and a student-friendly character. Therefore, a human character called Teacher Malik was chosen based on the demands of both of these characteristics. The human character was animated in a non-realistic 3-dimensional animated character (Fig. 4). From an animated character design aspect, a number of research have suggested several special designs for the face and eyes of the character aimed at avoiding the effect of the Uncanny Valley phenomenon on viewers, such as the size of the character's eyes being less than 50% of the actual size of a human eye, the position of the character's eyes, nose and mouth of an actual human and the character's face-body proportion should follow the design of an actual human, and the skin texture of the character should not be similar to that of a human [10]. These designs are considered important when developing an instructional-based animated character in order to the learning process. Besides, the display of the talking-head character is limited on the screen, which only portrays the head up to the shoulders [26]. Meanwhile, facial expression and lip synchronization are also important elements in developing the talking-head animation [27].
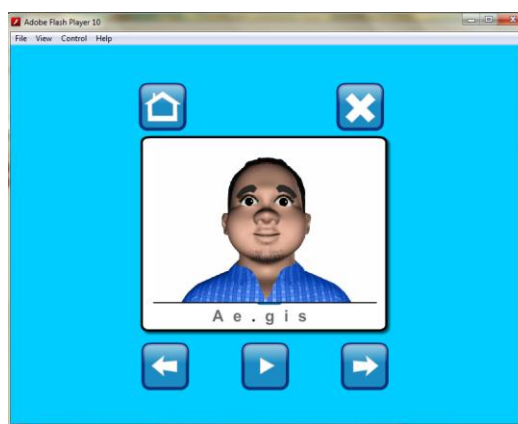


Fig. 4. An example of a non-realistic three-dimensional character.

## 5.5. Colour

Among the important components of graphics is colour [28]. Using too much colour, which is usually using more than five colours, is not suitable for designing a good animated presentation [28]. Therefore, the colour used in developing the talking-head animation courseware did not exceed five colours, as shown in Fig. 3. Blue was used as the courseware's background because the colour is characterised as being cool and suitable as a background colour [28]. Meanwhile, white was used in the animation block in order to attract

students' attention towards the animation. The colour of the font should be in contrast with the background in order to facilitate reading texts [28]. Hence, the font colours used in courseware were white and black.

## 6. Evaluation

The analysis process involving the 3D non-realistic talking-head animation courseware was carried out throughout the development process based on previous research, interviews and advice from several chosen experts. The analysis is not only limited to the physical appearance but also covers the overall performance, especially the applicability and user satisfaction when using the 3D non-realistic talking-head animation courseware. To evaluate the courseware's applicability and user satisfaction, a test was carried out as soon as the design and development process were completed. The questionnaire on applicability and user satisfaction (PSSUQ) was adapted to acquire the feedback from participants. The questionnaire consisted of 21 items that were divided into seven categories, such as design, function, convenience of use, ability to learn, satisfaction, future use, errors and reliability.

The questionnaire's items were 7-point scales with 1 for strongly agree, followed by next numbers in sequence with the end point 7 for strongly disagree. There were 50 participants in this study from the local Community College Certificate in Architecture Program. The study lasted 30 minutes after the students had explored the courseware. The teacher had explained the main function of the courseware before the students were allowed to explore the courseware at their own pace. After 20 minutes of exploration, the students were allowed to ask questions and have discussions during the session, which ended with the students answering the questionnaire. Generally, the alpha Cronbach value for the applicability and user satisfaction test using the PSSUQ questionnaire was .96. This value shows a high reliability for the items in the questionnaire.

*Design and Layout*: Overall, the students agreed with the design of the courseware interface (M = 3.34, SD = 1.40), the order of information given was clear and concise (M = 3.34, SD = 1.50), and the interface was pleasant (M = 3.56, SD = 1.32).

*Functionality:* The courseware had all the functions and capabilities anticipated by the student (M = 3.66, SD = 1.34), the information taken from the courseware was effective in assisting the students to complete their daily assignments (M = 3.40, SD = 1.31), and the students agreed that all the characteristics in the courseware function had functioned well (M = 3.32, SD = 1.59).

*Ease of use:* The students agreed that the courseware was convenient to use (M = 2.74, SD = 1.61), it was easy to find the required information (M = 3.26, SD = 1.48), the information given was precise (M = 3.14, SD = 1.55), and generally, the courseware was easy to use (M = 3, SD = 1.55).

*Learnability:* The students agreed that the courseware was convenient for learning (M = 2.98, SD = 1.55). They found that there was not much information that needed to be read before using the courseware (M = 3.28, SD = 1.56). They also agreed that the information given through the courseware was easily understood (M = 3.00, SD = 1.39).

*Satisfaction:* Generally, the students felt comfortable using the courseware (M = 3.20, SD = 1.68). They were happy and enjoyed exploring the courseware (M = 3.20, SD = 1.4), and overall, the students were satisfied with the courseware's development (M = 3.30, SD = 1.66).

*Future Use:* The students agreed that they could become productive at a faster pace when using the courseware (M = 3.36, SD = 1.35). They were confident that the courseware could enhance the pronunciation skills (M = 3.20, SD = 1.51), and based on the current experience, they were confident in frequently using the courseware again (M = 3.35, SD = 1.38).

*Errors and reliability:* The students agreed that every time they made a mistake in using this courseware, they could easily and quickly return to use the courseware again (M = 3.36, SD = 1.39). They agreed the courseware has been developed to give a clear message on the error and inform them on how to solve a problem (M = 3.74, SD = 1.58).

Overall, it shows that the students from the Community College agreed on the level of applicability and

satisfaction of using the 3D-NR talking-head animation courseware for learning to pronounce English words.

## 7. Conclusion

The 3D-NR talking-head animation is an instructional animation that is capable of enhancing communication skills through the enhancement of skilful pronunciation of a word correctly and accurately. This matter has been proven in several studies, whereby it has been found that the use of dynamic animation is better compared to conventional learning. Therefore, to ensure that students use the courseware with ease, applicability and user satisfaction tests were carried out and the analysis showed a positive effect from the student's perspective. This shows that the 3D-NR talking-head animation courseware meets the standards of an animated instruction, specifically for learning to pronounce English words.

## References

[1] Norabeerah, S., Halimah, B. Z., & Azlina, A. (2012). Penggunaan alat pengarangan multimedia dalam kalangan pendidik. *Jurnal Teknologi Pendidikan Malaysia*, *2(4)*, 5-14.

[2] Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education*, *21(1)*, 31–61.

[3] Muyassar, A. A. (2010) .The effects of using text and picture animation on promoting english learning among students of the 5th grade. Unpublished master's thesis. Hebron: Hebron University.

[4] Danihelka, J., Hak, R., Kencl, L., & Zara, J. (2011). 3D talking-head interface to voice-interactive services on mobile phones. *International Journal of Mobile Human Computer Interaction*, *3(2)*, 50-64.

[5] Wik, P., & Hjalmarrson, A. (2009). Embodied conversational agents in computer asistested languange learning. *Speech Communication*, *51(10)*, 1024-1037.

[6] Wang, H., Han, W., Soong, F., & Huo, Q. (2011). Text Driven 3d photo-realistic talking head. *Proceedings of the Interspeech* (pp. 307-3308)*.*

[7] Fraser, H. (2000). *Coordinating Improvements in Pronunciation Teaching for Adult Learners of English as a Second Language*. Canberra: DETYA.

[8] Moore, R. K. (2012). A Bayesian explanation of the 'Uncanny Valley' effect and related psychological phenomena. *Nature Scientific Reports*, *2(2)*, 10-23.

[9] Oke, O. K., & Alam, G. M. (2010). Comparative evaluation of the effectiveness of 2D and 3D visualizations in students' understanding of structures of organic molecules. *International Journal of the Physical Sciences*, *5(5),* 605–611.

[10] MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25(3)*, 695–710,

[11] MacDorman, K. F. (2005). Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it. *CogSci-2005 Workshop: Toward Social Mechanisms of Android Science*, 106-118.

[12] Tinwell, A., Grimshaw, M., Abdel Nabi, D., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, *27(2)*, 741-749.

[13] MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. *ICCS/CogSci-206 Long Symposium: Toward Social Mechanisms of Android Science*, 26-29.

[14] Mori, M. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine, 19(2)*, 98–100.

[15] Zamzuri, M. A., Farhana, M. I., & Syamsulaini, S. (2012). *Pembelajaran Berasaskan Web Isu Dan Trend*. Tanjung Malim: Penerbit Universiti Pendidikan Sultan Idris.

[16] Andrew, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development*, *3(4)*, 2-16.

[17] Gustafson, K. L., & Branch, R. M. (2002). What is instructional design? In Reiser, R, A., & Dempsey, J. V.

(Eds.), *Trends and Issues in Instructional Design and Technology*. Columbus: Prentice Hall.

[18] Kailani, I., & Muhammed, L. H. (2011). Pembangunan bahan e-pembelajaran berasaskan moodle bagi tajuk Ratios, rates and proportions 11 tingkatan tiga. *Journal of Educational Science and Mathematics*, *1(2)*, 35–43.

[19] Farag, M., & Shemy, N. (2011). Course delivery through the web: Effects of linear/nonlinear navigation and individual differences in online learning. *International Journal on e-Learning, 10(3)*, 243-271.

[20] Lee, H. L., & Olson, H. A. (2005). Hierarchical navigation: An exploration of yahoo directories. *Knowledge Organization*, *32(1)*, 20-24.

[21] Rozinah, J. (2005). *Multimedia Dalam Pendidikan*. Kuala Lumpur: Utusan.

[22] Zamzuri, M. A. (2012). Lima prinsip persembahan animasi dalam perisian pendidikan. *Jurnal Teknologi Pendidikan Malaysia*, *2(1)*, 15-22.

[23] Vaughan, T. (1998). *Multimedia: Making it work* (4th ed.). Barkley, KY: McGraw Hill.

[24] Peck, W. (2003). *Great web Typography*. Indiana: Wiley Publishing.

[25] Chandler, P. (1995). Is conventional computer instruction ineffective for learning? *Proceedings of Australian Computers in Education Conference* (pp. 322-332).

[26] Rieber, L. P., & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer generated animated visuals. *Journal of Computer Based Instruction, 18(3)*, 83-88.

[27] Dey, P., Maddock, S., & Nicolson, R. (2010). Evaluation of a viseme-driven talking head. *Proceedings of the EG UK Theory and Practice of Computer Graphics* (pp. 118-116).

[28] Fenrich, P. (1997). *Practical Guidelines for Creating Instructional Multimedia Applications*. Orlando: Dryden.

**Mohd Najib Hamdan** is currently a PhD student in the Faculty of Art, Computing and Creative Industry, Universiti Pendidikan Sultan Idris, Malaysia. He has his bachelor degree in information technology (software engineering) and the master in education from Universiti Tun Hussein Onn Malaysia. His research interest is instructional technology.

**Ahmad Zamzuri Mohamad Ali** is an associate professor of multimedia in the Faculty of Art, Computing and Creative Industry, Universiti Pendidikan Sultan Idris, Malaysia. He has a bachelor degree in electrical engineering from Universiti Teknologi Malaysia, a master in education from Universiti Teknologi Malaysia and the PhD in multimedia design from Universiti Sains Malaysia. He has taught both face-to-face and online classes in higher education for over 15 years. His research and publication interests are multimedia design, instructional technology, ICT in education and open source in education.