# The Cognitive Effects of Different Feedback Modalities in Virtual Reality Learning

## Sze Cha TAN<sup>a</sup> & Chwen Jen CHEN<sup>b\*</sup>

<sup>ab</sup>Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak \*cjchen@fcs.unimas.my

Abstract: Feedback is an important feature of a virtual reality (VR) learning system as appropriate feedback increases learners' motivation as well as interaction. This study aims to investigate the cognitive effects of using VR-based learning systems with different feedback modalities. It employs a pretest-posttest quasi-experimental design to measure the cognitive effects of VR learning systems that use on-screen text, narration, as well as both on-screen text and narration, respectively and an additional non-VR environment, which uses paper-based reference material, to serve as a control of the study. The study reveals that the differences in feedback modality, focusing on narration and on-screen text, do not significantly affect cognitive gain in a VR learning system. In addition, the significant positive effects of the VR-based learning system when compared with the non-VR method, provides another evidence of the potentials of VR technology for instructional use. This paper discusses the findings based on existing learning theories and principles, and concludes with a design implication of VR learning systems that incorporate feedback.

Keywords: virtual reality, learning, feedback

#### Introduction

Feedback is a crucial component of any learning process [1], [6], [8], & [17]. It informs learners about the consequences of their actions and motivates them to further interact with a system [25]. Interaction with a system can only be considered as a contributor to the learning process if the learner gets feedback on what his or her action will result in. According to Mulder, Elgar, and Brady in [16], effective learning occurs when learners obtain timely and detailed feedback on their performance from their instructors as well as peers.

## 1. Modality Effects

Past research in multimedia learning demonstrated a modality effect in which students who studied from pictures and spoken words outperformed students who studied the same pictures with text [14]. According to Mousavi, Low and Sweller in [15], students learn better when the verbal information is presented auditorily as speech rather than visually as on-screen text for both concurrent and sequential presentations. Mayer, Heiser and Lonn in [11] investigate the redundancy effect consistent with a dual-channel theory of multimedia learning in which adding on-screen text can overload the visual information-processing channel, causing learners to split their visual attention between two sources. One of the important findings is that audio feedback is perceived to be one of the most important features that engender a sense of presence [5]. Marila and Ronkainen in [9] also reveals that auditory feedback enables faster response times than visual feedback.

Mayer and Moreno in [12] carried out a research in which they requested students to view an animation depicting a complex multimedia system, either along with concurrent narration or along with concurrent on-screen text. Their study shows that students who learned with concurrent narration and animation outperformed those who learned with concurrent ICCE2010 | 591

on-screen text and animation. The cognitive capacity increases when information is in mixed (auditory and visual mode) rather than in a single mode [15]. The opportunities for active cognitive processing are reduced when working memory becomes overloaded [14]. Mayer and Moreno in [12] suggests that animations should be accompanied by narration rather than by on-screen text when designing multimedia presentations where working memory is less likely to become overloaded.

Many studies found that students learn better when material is auditorily-presented than visually-presented. Auditory feedback is shown to bring about more positive effects compared with other types of feedback. However, it is important to note that these studies are mostly based on multimedia systems. Indeed, there is still a lack of similar research for virtual reality (VR) learning system. Feedback is an important feature of a VR learning system. According to Jung in [7], a VR learning system with appropriate feedback increases the learners' motivation and interaction. Hence, this study takes the effort to investigate the cognitive effects of using feedback in the form of narration, on-screen text, or both narration and on-screen text in a VR learning system.

## 2. Aim

The study aims to investigate the cognitive effects of using VR-based learning systems that use on-screen text (VR-T), narration (VR-N), as well as both on-screen text and narration (VR-N&T), respectively and an additional non-VR environment (Non-VR), which uses paper-based reference material, to serve as a control of the study.

## 2.1 Research Design

A learning mode refers to each of the VR-based learning systems, including the non-VR environment. The study employs a pretest-posttest quasi-experimental design to measure the cognitive effects of each learning mode. Figure 1 depicts this research design.



Figure 1: Research design

## 2.2 Research Question and Hypothesis

This study attempts to answer the following research question:

Is there a difference in the posttest score of the VR-based test between the four learning modes (VR-T, VR-N, VR-N&T, and Non-VR)?

The statement of the null hypothesis that corresponds to the above-stated research question is as follows.

There is no significant difference in the posttest score of the VR-based test between learners exposed to the VR-T mode, learners exposed to the VR-N mode, learners exposed to the VR-N&T mode, and learners of the control group.

The outcomes depict the cognitive effects of VR-based learning systems with different feedback modalities. The understanding of such effects will help to add to the suggestions of how to effectively utilise the VR capabilities to support the desired learning outcomes.

# 3. Methods

## 3.1 VR Learning System

The VR learning system that served as the treatment for the experimental groups was meant to provide novice car drivers, the targeted learners, with the knowledge on basic car maintenance procedures. These procedures include pumping car tyre, changing punctured tyre, adding coolant, adding engine oil, as well as changing engine oil and oil filter. Figure 2 shows a screenshot of the learning system, depicting how its design is guided by the instructional design model for VR-based learning environment as proposed in [2]. As this paper focuses on the findings of the investigation on the effects of different feedback modalities, it does not provide a detail explanation on how the system was designed based on this model.



Figure 2: A screenshot of the learning system, depicting how its design is guided by the instructional design model

For learning to occur, a system should allow feedback about the learners' performance on the tasks [15]. If feedback can be built in automatically to the practice routine in some way then it is unnecessary to have an instructor to permanently present as the required feedback can be contained within the system. Hence, by incorporating feedback into the VR-based learning system of this study, it replaces some of the roles that are conventionally played by the instructors of the driving school.

# 3.2 Population and Sample

The current law in Malaysia allows any person who is 17 years old and above to undergo the 6-hour Class Briefing/Basic Start session in a driving institute or academy. Although the age of these candidates may vary greatly, majority of them are from the younger group, those who are just above the eligible age [2]. Therefore, this majority forms the targeted population of this study.

The accessible population for this study encompassed first-year students of any colleges or universities that are well equipped with multimedia computer laboratories in Kuching, Sarawak. College or universities students are chosen, as most of them are computer and information technology literate. Besides, they are chosen instead of general public, in order to obtain better-controlled samples.

In this study, students of two private educational institutions in Kuching, a college and a university, formed the accessible population. This also implies the findings of this study can only be strictly applied to this group of population. A total of 200 university and college students participated in this experiment. There were five intact classes with 120 first year students from one of the institutions and four intact classes with 80 first year students from another institution participated in this experiment. However, only a total of 153 students were taken into consideration in the analyses as some of them were absent during either the

pretest or posttest session. The mean age of the learners was 19.15 years old. Table 1 shows the distribution of learners across learning modes.

Table 1: Learners distribution across learning modes			
Learning mode	Number of learners		
VR-T	39		
VR-N	38		
VR-N&T	42		
Non-VR	34		

#### 3.3 Instruments – VR-based Test (Pretest and Posttest)

The VR-based test (pretest and posttest) is the instrument for this study. Both the VR-based pretest and VR-based posttest are computer-based tests. Each test comprises fifteen multiple-choice questions. The fifteen multiple-choice questions assess the learners' ability to identify the possible missing or incorrect car maintenance procedure performed in the simulations. Each of the fifteen multiple-choice questions is displayed through the web interface showing the maintenance process of a three-dimensional virtual car with possible missing or incorrect step. Both pretest and posttest are similar in content but the order of the questions is different to avoid the set response effect. The total score of each test will be 15. For each question, participants received a score of either 1 (correct answer) or 0 (incorrect answer), and a total score ranging from 0 to 15. This score is multiplied by 100 to convert it to percentage. The Cronbach's alpha reliability coefficient of the pretest was 0.742, which depicted the test items were satisfactorily reliable.

#### 3.4 Procedures

Prior to this study, permissions were obtained from the two institutions for conducting the pilot study as well as the experimental study. In order to prevent any experimental bias, lecturers from the two institutions were asked to conduct the pilot and experimental sessions. The researcher was only present to provide navigation training to the lecturers before the pilot and experimental session were conducted. The lecturers were given precise procedure on how to carry out the sessions smoothly. Students who were not selected as the sample to participate in the experiment were asked to provide assistance during the pretest and posttest sessions.

#### 4. Results

In this study, the pretest score served as the covariate and this pretest was administered before the participants of the four groups underwent their respective learning session to prevent the pretest score being influenced by any of the treatments. Preliminary checks were conducted to ensure the appropriateness of the use of pretest score as the covariate.

## 4.1 Testing of Hypothesis

One-way analysis of covariance (ANCOVA) was used to measure and analyse the collected data. After adjusting the pretest scores, there was a significant difference between the four learning modes on the posttest scores, F(3, 148) = 54.457, p = 0.000. The effect size, calculated using  $\eta^2$ , was 0.525, which would be considered as a large effect size. There was

also a strong relationship between the pretest scores and the posttest scores, as indicated by the  $\eta^2$  of 0.377.

The means, standard deviations, adjusted means, and standard error of the dependent variable (posttest score) by the learning mode were measured in Table 2. The standard deviations were used to measure the variability of the posttest scores. The VR-N&T mode with smallest standard deviation value (SD = 7.67079), shows that there is less heterogeneity with these groups when compared with the VR-N mode (SD = 9.93678), the VR-T mode (SD = 10.63806) and Non-VR mode with the largest standard deviation value (SD = 12.47448). The VR-N&T mode had the largest adjusted mean (adjusted M = 90.685), followed by the VR-N mode and VR-T mode (adjusted M = 87.227 and adjusted M = 86.450 respectively). The Non-VR mode had the smallest adjusted mean (adjusted M = 68.184).

	score by learning mode			
Learning mode	Posttest score (%)			
	М	SD	Adjusted M	SE
VR-N&T, $N = 42$	90.4755	7.67079	90.685 <sup><i>a</i></sup>	1.247
VR-T, N = 39	87.1787	10.63806	86.450 <sup><i>a</i></sup>	1.296
VR-N, N = 38	88.7711	9.93678	87.227 <sup><i>a</i></sup>	1.321
Non-VR, $N = 34$	65.8824	12.47448	68.184 <sup>a</sup>	1.407

 Table 2: Means, standard deviations, adjusted means, and standard errors of posttest

 score by learning mode

**Note:** <sup>*a*</sup> Evaluated at covariate appeared in the model: pretest = 64.0956

#### 4.2 Pair Wise Comparisons for One-Way ANCOVA

The follow-up post-hoc pair wise comparisons were conducted when the result of the one-way ANCOVA was found statistically significant. In this study, the Holm's sequential Bonferroni procedure was used to control for Type I error across the six pair wise comparisons.

Three comparisons were found significant: the comparison between the VR-T mode and the Non-VR mode (p of 0.000 is less than 0.0083), the comparison between the VR-N mode and the Non-VR mode (p of 0.000 is less than 0.0100), and the comparison between the VR-N&T mode and the Non-VR mode (p of 0.000 is less than 0.0125). Whereas, another three comparisons were found not significant: the comparison between the VR-T mode and the VR-N&T mode (p of 0.0200 is not less than 0.0167), the comparison between the VR-N mode and the VR-N&T mode (p of 0.0200 is not less than 0.0167), the comparison between the VR-N mode and the VR-N&T mode (p of 0.059 is not less than 0.025), and the comparison between the VR-T mode and the VR-N mode (p of 0.674 is not less than 0.05).

The analysis revealed that there were significant differences in the adjusted means between: the VR-T mode and the Non-VR mode, between the VR-N mode and the Non-VR mode, and between the VR-N&T mode and the Non-VR mode. However, the analysis revealed that there were no significant differences in the adjusted means between: the VR-T mode and the VR-N&T mode, the VR-N mode and the VR-N&T mode, and the VR-N mode and the VR-N mode and the VR-N mode.

#### 5. Discussion

The insignificant difference between the posttest results of the learners exposed to VR-T, VR-N, and VR-N&T modes can be explained according to the Cone of Learning by Dale (1969) as well as the limited-capacity assumption and dual-coding assumption of Mayer's (2001) cognitive theory of multimedia learning.

#### 5.1 Cone of Learning

Dale's Cone of Learning in [4] is based on the relationships of various educational experiences to reality (real life). At the upper part of the cone which considers experiences that are far from reality, learners tend to remember only 10% of what they read. Learners, who learn via hearing words, tend to remember 20% of what they hear while those who learn by looking at pictures tend to remember 30% of what they see. As for learners who learn by hearing and seeing images at the same time, they tend to remember more, 50% of what they hear and see. Conversely, the lower part of the cone considers educational experiences that close to real and everyday life. Such mimic to real-world experiences stimulate the use of all senses, which include seeing, smelling, hearing, touching and moving. According to Dale in [4], better learning will occur when more sensory channels are involved in interacting with an educational resource. Hence, the lower part of the cone stresses the importance of learning by doing and suggests that the learners tend to remember 90% of what they both say and do.

All the three learning modes require the learners to experience and perform the maintenance procedures. According to the Cone of Learning, such active learning by doing tend to help the learners to remember up to 90% of the learning content while they interact with the virtual system. The Cone of Learning also explains the use of different feedback modalities tend to produce different amount of recalling ability. Learners exposed to the VR-T, VR-N and VR-N&T mode tend to remember 10%, 20% and 50% respectively of the feedback. This may explain the slightly higher posttest scores achieved by the learners of VR-N&T mode than the learners exposed to either VR-T or VR-N mode. Nevertheless, such differences which are due to different types of feedback used are insignificant because feedback only forms a portion of the overall learning process, while a major part of the learning process focuses on the manipulation of the different parts of the virtual car. This helps to explain the insignificant difference between the three learning modes.

#### 5.2 Cognitive theory of multimedia learning

The three fundamental assumptions underlying the cognitive theory of multimedia are dual channels, limited capacity, and active processing [10]. This theory assumes that the human information processing system includes two channels; a) visual or pictorial, and b) auditory or verbal processing. Each channel has limited capacity for processing, and that active learning entails carrying out a coordinated set of cognitive processes during learning.

The limited capacity assumption of this theory stresses that human can pay attention to only a few pieces of information in each channel at a time, hence, it is important not to overload the working memory during the learning process [22]. According to Cooper in [3], when the intrinsic cognitive load is high (difficult domain concepts or knowledge) and the extraneous cognitive load is high, then total cognitive load will exceed mental resources and learning may fail to occur. In order to reduce the total cognitive load to within the bounds of mental resources, the level of extraneous cognitive load must be modified by changing the instructional materials presented to learners [3]. The dual-channel assumption of this theory posits that humans possess separate information channels for visually presented material and auditorily presented material.

Mayer and Moreno [12] in their research on multimedia have shown that students who learned with concurrent narration and animation outperformed those who learned with concurrent on-screen text and animation. In their research context, students who experienced concurrent narration and animation used the auditory channel to process the information from narration and the visual channel for animation. On the other hand, students who experienced concurrent on-screen text and animation used only the visual

channel to process information from both the on-screen text as well as the animation. This has somehow overloaded to visual channel, which consequently resulted in poorer performance.

In this study, the use of narration within the dynamic virtual environment, as in VR-N, involves a learner's auditory channel to process the narration and visual channel to process information from the real-time three-dimensional graphical representation of the virtual environment. The use of on-screen text in VR-T evokes the visual channel of a learner. This visual channel is not overloaded with the information from the virtual environment because the dynamicity of the virtual environment in VR-T is halted when the on-screen text is displayed. In other words, the learner focuses only on the on-screen text when such feedback appears. In the case of VR-N&T, the virtual environment is once again halted when a feedback appears. A learner will only uses the auditory channel to process the narration and visual channel to process the on-screen text. Hence, in all these learning modes, neither visual nor the auditory channel is overloaded. This helps to explain the insignificant difference in the cognitive gain of the three learning modes.

Based on the dual-channel assumption, the use of both channels is supposed to produce better learning as it creates two routes to retrieve information from memory [19]. However, Mayer in [10] asserts that when learners are able to allocate sufficient cognitive resources to a task, it is possible for information originally presented to one channel to be represented in the other channel. In VR-T mode, when information was presented to the learners' eyes (on-screen text), they started to process the information in the visual channel. In VR-N mode, when information was presented to the learners' ears (narration), learners started to process the information in the auditory channel. Due to the simplicity of the feedback message as explained earlier, it is possible that adequate cognitive resources are allocated for the cognitive processing of it. For example, when the on-screen feedback text of VR-T mode, such as "Open the car hood", is initially presented to the eyes, learners may mentally convert the open-the-car-hood images into sound, which is processed through the auditory channel. Similarly in the VR-N mode, when the narration describing the event such as "The engine oil is too dirty" is initially presented to the ears, learners may also form a corresponding mental image that is processed in the visual channel. Therefore, the VR-T and VR-N modes did not seem to produce significant differences when compared with the VR-N&T mode although originally the text or audio message gets into a single information processing channel. The cross-channel representations of the same stimulus play a vital part in Paivio's dual-coding theory [18].

## 6. Conclusion

The study reveals that the differences in feedback modality, focusing on narration and on-screen text, do not significantly affect cognitive gain in a VR learning system. Hence, the instructional designer of a VR learning system may choose to use any of these feedback modalities without jeopardizing its effectiveness, at least in a learning context that is similar to the one used in this study. In addition, the significant positive effects of the VR-based learning system when compared with the non-VR method, provides another evidence of the potentials of VR technology for instructional use. This VR-based learning system provides new learning opportunities by introducing learning activities that make visible concepts and relationships that are not easily grasped or visualized by learners when relying only on the conventional method.

The study only involves college and university students. As for the future, the work can be extended to include other groups of learners to improve the generalisability of the findings. More research studies are recommended to verify if the similar results will be attained if the VR-based learning system focuses on abstract concepts rather than concrete tasks. The

VR-based learning system in this study only focuses on the simulation of a real-world system, the real-world car maintenance procedures, which correspond to a concrete task. Future initiatives may also include investigations into other VR characteristics such as the control over view position and direction, representation fidelity, ability to manipulate virtual objects, user interface as well as navigation in the virtual world. This may help in generating more useful principles for designing effective VR-based learning systems.

#### Acknowledgements

We would like to thank Universiti Malaysia Sarawak for providing us necessary facilities to conduct this study and financial support to attend this conference.

#### References

- [1] Brown, A. (1997) Designing for learning: What are the essential features of an effective online course? *Australian Journal of Educational Technology*, *13*(2), 115-126.
- [2] Chen, C.J., Toh, S.C. & Wan, M.F. (2004). The theoretical model for designing desktop virtual reality-based learning environments. *Journal of Interactive Learning Research*, 15(2), 147-167.
- [3] Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW. Retrieved from http://education.arts.unsw.edu.au/staff/sweller/clt/
- [4] Dale, E. (1969). Audio-visual methods in teaching (3<sup>rd</sup> ed.). New York: Holt, Rinehart & Winston.
- [5] Jelfs, A., & Whitelock, D. (2000). The notion of presence in virtual learning environments: What makes the environment 'real'? *British Journal of Educational Technology*, *31*(2), 145 152.
- [6] Jonassen, D.H. & Hannum, W.H. (1987). Research-based principles for designing computer software. *Educational Technology*, 27(12), 7-14.
- [7] Jung, H. J. (2002). Virtual reality for ESL students. *The Internet TESL Journal*, 8(10). Retrieved from http://iteslj.org/Articles/Jung-VR.html
- [8] Laurillard, D. (1993). *Rethinking university teaching: A framework for effective use of educational technology*. London: Routledge.
- [9] Marila, J., & Ronkainen, S. (2003). Time-out in mobile text input. The effects of learning and feedback. In L. Chitaro (Ed.), *The 5<sup>th</sup> International Symposium Mobile HCI 2003 Proceedings*. Berlin /
- [10] Heidelberg / New York: Springer-Verlag.
- [11] Mayer, R. E. (2001). *Multimedia Learning*. Cambridge: Cambridge University Press.
- [12] Mayer, R. E., Heiser, J., and Lonn, S. (2001). Cognitive constraints on multimedia learning: When
- [13] presenting more material results in less understanding. *Journal of Educational Psychology*, 93(1), 187-198.
- [14] Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 312-320.
- [15] Mayes, J.T. (2002). The technology of learning in a social world. In R. Harrison, F.Reeve, A. Hanson, J.Clarke (Eds.), Supporting lifelong learning: Perspectives on learning Vol. 1, Routledge: London
- [16] Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- [17] Mousavi, S. Y., Low, R., & Sweller, J. (1995) Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.
- [18] Mulder, R.A., Elgar, M.A., & Brady, D. (2005). "APRES": Electronically managed student feedback via peer review. *Proceedings of the Blended Learning in Science Teaching and Learning Symposium*. (pp. 2-6).
- [19] Norman, G.A., & H.G. Schmidt. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine*. 67, 557–565.
- [20] Paivio, A. (1986). *Mental representations*. New York: Oxford University Press.
- [21] Rieber, L. P. (2005). Multimedia learning in games, simulations, and micro worlds. In: R.E. Mayer (Ed.), *Cambridge handbook of multimedia learning*. (pp. 549-568). New York: Cambridge University Press.
- [22] Sweller, J. (1999). Instructional design in technical areas. Camberwell, Australia: ACER.
- [23] Vasilyeva, E., Pechenizkiy, M., Puuronen, S. (2006). The challenge of feedback personalization to learning styles in a web-based learning system. *Proceedings 6th IEEE International Conference on Advanced Learning Technologies*. (pp. 1143-1144). IEEE Computer Society.