Learning with desktop virtual reality: Low spatial ability learners are more positively affected

Elinda Ai-Lim Lee, Kok Wai Wong

PII: S0360-1315(14)00161-4
DOI: 10.1016/j.compedu.2014.07.010
Reference: CAE 2673

To appear in: *Computers & Education*

Received Date: 17 January 2014
Revised Date: 21 July 2014
Accepted Date: 22 July 2014

Please cite this article as: Ai-Lim LeeE. & WongK.W., Learning with desktop virtual reality: Low spatial ability learners are more positively affected, *Computers & Education* (2014), doi: 10.1016/j.compedu.2014.07.010.

Manuscript Title: Learning with desktop virtual reality: Low spatial ability learners are more positively affected

Elinda Ai-Lim Lee[a] (Corresponding author), Kok Wai Wong[b]

[a] School of Occupational Therapy and Social Work, Curtin University
GPO Box U1987, Perth, Western Australia 6845
Email: Elinda.Lee@curtin.edu.au
Tel: +618-9266 1790
Fax: +618-9266 3636

[b]Murdoch University
School of Engineering and Information Technology
South St, Murdoch
Western Australia 6150
Email: k.wong@murdoch.edu.au
Tel: +618 9360 6100
Fax: +618 9360 2941

Title: Learning with desktop virtual reality: Low spatial ability learners are more positively affected

Abstract

This study aims to verify the learning effectiveness of a desktop virtual reality (VR)-based learning environment, and to investigate the effects of desktop VR-based learning environment on learners with different spatial abilities. The learning outcome was measured cognitively through academic performance. A quasi pretest-posttest experimental design was employed for this study. A total of 431 high school students from four randomly selected schools participated in this study where they were randomly assigned to either experimental or control groups based on intact classes. Findings indicate a significant difference in the performance achievement between the two groups with students performed better using desktop virtual reality. A possible explanation is that the desktop virtual reality instructional intervention has helped to reduce extraneous cognitive load and engages learners in active processing of instructional material to increase germane cognitive load. A significant interaction effect was found between the learning mode and spatial ability with regard to the performance achievement. Further analysis shows a significant difference in the performance of low spatial ability learners in the experimental and control groups, but no statistically significant difference in the performance for high spatial learners in both groups. The results signify that low spatial ability learners' performance, compared with high spatial ability learners, appeared to be more positively affected by the desktop VR-based learning environment which is supported by the ability-as-compensator hypothesis, and can be explained by the cognitive load theory.

Keywords Desktop virtual reality, performance achievement, spatial ability, learning, high school students

1. Introduction

There is a growing trend to use virtual reality (VR)-based learning in schools and colleges (Huang, Rauch, & Liaw, 2010; Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014; Mikropoulos & Natsis, 2011; Stull, 2009). De Jong, Linn and Zacharia (2013) has elaborated the use and advantages of virtual laboratories in science and engineering education in the recent decade. Nevertheless, research finding are mixed with regard to the learning effectiveness of VR-based learning. Positive research outcomes have been reported with VR-based learning such as better performance in business knowledge application (Cheng & Wang, 2011); improved anatomy learning (Petersson, Sinkvist, Wang, & Smedby, 2009); greater efficiency in matching diagrams and models (Stull, Barrett, & Hegarty, 2013); improved calligraphic writing skills (Wu, Yuan, Zhou, & Cai, 2013); improved spatial thinking (Cohen & Hegarty, 2014; Dünser, Steinbügl, Kaufmann, & Glück, 2006; Hauptman, 2010); enhanced spatial abilities for "sensing" and kinesthetic learning style learners (Hauptman & Cohen, 2011) and for low visual spatial ability learners (Meijer & van den Broek, 2010); and the ability to accommodate learners with different learning styles in cognitive outcomes (Chen, Toh, & Wan, 2005; Lee, Wong, & Fung, 2010b) and affective outcomes (Lee et al., 2010b)

On the other hand, Urhahne, Nick, and Schanze (2009) found no difference in understanding chemical structures and their properties for freshman students using 3-D simulations and two-dimensional (2-D) images. Similarly, in the study of Merchant, Goetz, Keeney-Kennicutt, Cifuentes, Kwokt, and Davis (2013), the hypothesis that 3-D virtual environment can enhance chemistry learning among undergraduate students was not supported. Besides, students did not demonstrate greater understanding of the genetics concepts in the study of Annetta, Minogue, Holmes, and Cheng (2009). Therefore, VR might not work for all kinds of learning. Learner characteristics or individual differences can account for different learning results in VR-based learning environment (Chen 2006; Hauptman & Cohen, 2011). In recent years, there is more focus on the role of learner characteristics or individual differences on learning with visual representations (Höffler & Leutner, 2011). The importance of considering individual differences in visual representations is also emphasized by Meijer and van den Broek (2010). The effects of learner characteristics on learning outcomes would enable instructor to adapt the nature of instruction to accommodate individual differences to improve learning outcomes.

This study therefore aims to verify the learning effectiveness of a desktop VR-based learning environment in biology education, and to investigate the effects of VR-based learning environment on learners with different spatial abilities. We intend to provide an answer to these questions: (1) Is there any difference in the learners' performance achievement between a desktop VR-based learning environment and a conventional classroom learning practice? (2) What is the interaction between learners' spatial ability with the learning environment with regard to their performance achievement?

2. What is virtual reality?

VR is a way of stimulating or replicating an environment that can be explored and interacted with by a person (Ausburn & Ausburn, 2004; Chuah, Chen, & Teh, 2008; Inoue, 2007). VR computer simulations could take many forms, ranging from computer renderings of 3-D geometric shapes on a desktop computer to highly interactive, fully immersive multisensory environment in laboratory (Ausburn & Ausburn, 2004; Mikropoulos & Natsis, 2011; Strangman & Hall, 2003). Unlike static picture, the VR computer image is dynamic, a mimic of a real object and can be rotated to different orientations with a handheld device (Stull, 2009). Education has moved to the use interactive technologies to help impart knowledge and understanding as an alternative to books, pencils and pens (e.g. Cheng & Wang, 2011; Aoki, Oman, Buckland & Natapoff, 2008; Kebrtichi, Hirumi & Bai, 2010).

There are basically two types of VR: Immersive VR and Non-immersive VR. Due to the technological advancements, today's VR system can run on a relatively cheap system such as desktop personal computer, which is known as "non-immersive" or "desktop" VR (Chen, Toh, & Wan, 2004; Lee, Wong, & Fung, 2009; Merchant et al., 2014; Strangman & Hall, 2003). Non-immersive VR or desktop VR is a 3-D image that generated in a multimedia environment on a personal computer, which can be explored interactively by using keyboard, mouse, joystick or touch screen, headphones, shutter glasses, and data gloves (Chen et al., 2004; Gazit, Yair, & Chen, 2006; Strangman & Hall, 2003). The desktop VR-based learning environment could be games, simulations or virtual worlds. The advancement of web technologies has enabled multiple users to work collaboratively in a virtual environment such as Second Life®. Though desktop VR is considered less immersive; however, Dalgarno, Hedberg and Harper (2002) argue that "the sense of presence or immersion in a virtual environment is induced by the representational fidelity and the high degree of interaction and control of user, rather than just a unique attribute of the environment". Immersive VR environments are presented on multiple, room-size screen or through a stereoscopic, head-mounted display unit (Chen et al., 2004; Dalgarno et al., 2002; Strangman & Hall, 2003). Lee & Wong (2008) has articulated the three levels of immersive VR classified by Allen, Austin, Beach, Bergstrom, Exon, Fabri et al. (2002): partially or semi immersive VR; fully immersive VR; and augmented reality or mixed reality. Due to the high cost of immersive VR systems and the inherent problems associated with them such as simulator sickness, desktop VR provides an alternative to immersive VR systems because it retains the benefits of real time visualization and interaction within a virtual world (Chen et al., 2004; Chuah et al., 2008; Huang et al., 2010; Merchant et al., 2014; Merchant, Goetz, Keeney-Kennicutt, Kwok, Cifuentes, & Davis, 2012).

3. Theoretical background

3.1. Aptitude-by-treatment interaction

Aptitude-by-treatment interaction (ATI) research refers to the concept that instructional strategies are more effective when they are adapted to the specific abilities and/or attributes of the learners (Fletcher & Tobias, 2005, p. 130; Plass, Kalyuga, & Leutner, 2010). Ausburn and Ausburn (2004) have called for the application of the ATI model in new studies in VR in education because the ATI model is more multi-factor in concept which involves not only independent and dependent variables but moderating variable as well. In this model, the interest is not on the effect of an instructional method, if it works or is better, but on the interactions between various instructional methods and learners' aptitudes or characteristics. Interaction between aptitude and treatment occurs when the effect of treatment differs depending on the level of aptitude measure. Such a research model will enlighten educators for what purposes and for whom an instructional method may be effective (Ausburn & Ausburn, 2004). There is limited research that investigates the statistical effect of ATI between instructions and spatial ability (Wang, Chang, & Li, 2007).

3.2. Spatial Ability and VR

Spatial ability refers to a group of cognitive functions and aptitudes that is crucial in solving problems that involve manipulating and processing visuo-spatial information (Lajoie, 2008). Spatial visualization ability is a measure of the ability to mentally restructure or manipulate the components of visual stimulus and involves recognizing, retaining, recalling configurations when the figure or parts are moved (McGee, 1979). It is believed that spatial visualization ability is the primary cognitive factor that causes the differences in performance and has an impact on comprehension of 3-D computer visualization (Huk, 2006). Students with different spatial ability will benefit differently when learning with interactive 3-D animation or simulations (Höffler & Leutner, 2011; Huk, 2006) which depends on their ability to extract relevant information and then to reconstruct or incorporate the information into their existing mental models.

The study of Merchant et al. (2012) reported that spatial orientation mediates the relationships between 3D virtual learning environment features and chemistry learning outcomes. Lee, Wong, and Fung (2010a) have also found that control and active learning in a VR learning environment is a more concern factor for the high spatial ability learners. High spatial ability learners are more likely to perform better, with higher level of perceived learning and satisfaction if control and active learning is provided (Lee et al., 2010a). Thus, VR technology might not benefit everyone equally with regard to spatial ability. This could be explained by the hypotheses as follows.

3.3. Interaction effect of spatial ability and learning design

When comparing the performance of high and low spatial ability learners between a desktop VR learning environment and a learning environment without interactive 3-D visualization, the ability-as-compensator hypothesis posits that low spatial ability learners benefit most from the VR-based learning environment because they have difficulty to mentally reconstruct their own visualization (Huk, 2006; Mayer, 2001). However, high spatial ability learners do not gain particular benefit because they are able to build their own visual representation based on static images and words alone (Mayer, 2001). This implies that high spatial ability compensates the non-interactive 3-D learning environment. On the other hand, based on the ability-as-enhancer hypothesis, high spatial ability should benefit particularly from the VR-based learning environment as they have enough cognitive capability left for mental model construction (Huk, 2006; Mayer, 2001). As for the low spatial ability learners, they do not gain particular benefit because the learning in the VR-based learning environment requires cognitive capability that exceeds the available memory resources (Mayer, 2001).

3.4. Cognitive Load Theory

Cognitive load theory (CLT) is an instructional theory based on our knowledge of human cognitive architecture that specifically addresses the limitations of working memory (Sweller, 2005, p. 28; 2010a, p. 29). CLT has become one of the fundamental theories used to describe the cognitive processes in learning with new technologies since the early 2000s (Brunken, Plass, & Leutner, 2003; Mayer, 2001). It has been increasingly used to inform the instructional design and to predict the learning effectiveness with new technologies (Brunken et al., 2003). Cognitive load refers to the total amount of mental activity imposed on working memory when processing information (Cooper, 1998; Sweller, 2005). The process of learning requires working memory to be actively engaged in comprehension of instruction material to encode to-be-learned information for appropriate schema construction that will be stored in long-term memory (Cooper, 1998). The capacity of working memory is limited, thus cognitive load theory asserts that learning is inhibited when the working memory capacity is exceeded by the total cognitive load in a learning task (Cooper, 1998; De Jong, 2010).

There are three types of cognitive load: intrinsic, extraneous and germane. Intrinsic cognitive load refers to the difficulty of the content and cannot be changed by instructional design and treatments. This is because it is determined by the interaction between the nature of materials being learned and the expertise of the learner (Sweller, 2010b; van Merriënboer & Sweller, 2005). Extraneous cognitive load is the load that does not directly contribute to learning (i.e., schema construction), which is evoked by the instructional method and it can be altered by instructional interventions (van Merriënboer & Sweller, 2005). Germane cognitive load is the load that is necessary for learning and it results from the way information is presented and the circumstances in which it is presented (Khalil, Paas, Johnson, & Payer, 2005). These three cognitive loads are additive and the sum should not exceed the memory resources available. For learning to be effective, activities and representations that maximize germane load should be provided while extraneous load should be minimized to ensure that the total cognitive load is within the memory resources available (Sweller, 2010b). If the learning content is difficult (high intrinsic cognitive load) and if the strategy to present the information has created a high extraneous cognitive load, then the total cognitive load may exceed the available memory resources. Consequently, learning will be impeded and may fail to occur (Sweller, 2010b). The extent to which instructional features contribute to extraneous load or germane load depend on the individual learner and the extent to which the individual learner experiences intrinsic load (Leppink, Paas, Van der Vleuten, Van Gog, & Van Merrienboer, 2013, p. 1058) .

4. Research objectives and hypotheses

The learning outcome was measured cognitively through academic performance, which was measured through a summative assessment. A VR-based frog dissection software program designed for biology education, V-Frog^{TM}, was used as the VR learning material (Tactus Techologies, 2007). The specific objectives of this research are:

1. To verify the learning effectiveness of a desktop VR-based learning environment (VR mode) against a conventional classroom learning practice (Non VR mode).
2. To investigate the interaction effect between the learners' spatial ability and the learning mode with respect to performance achievement.

In pursuance of the research objectives, the following hypotheses were formulated for testing.

H1: There is no significant difference in the performance achievement between students in the VR-based learning environment (VR mode) and the conventional classroom learning practice (Non VR mode).

H2: There is no interaction effect between the learners' spatial ability and the learning mode, related to the performance achievement.

5. Methodology

5.1. Research Design

This study used a pretest-posttest quasi-experimental design. This design was employed because classes could not be reorganized as the participants could only participate in the experiment at certain time of the school hours based on their class timetable. Therefore, intact classes were randomly assigned to either the experimental group or control group. The experimental group underwent a self-directed lesson on frog anatomy with desktop VR software, V-Frog^{TM} while the control group followed a similar lesson using the conventional classroom learning method with PowerPoint slides. The conventional classroom learning method was conducted by the class biology teacher. Both groups were given a pretest, posttest and spatial ability test. The experimental process is as shown in Fig. 1. In order to minimize the learning content differences between the two learning modes, the PowerPoint slides have similar texts and coloured pictures as in the desktop VR software, but with no animations. The use of PowerPoint slides could also help to minimize the teaching capability differences of teachers from different classes and different schools in the control groups because the contents and the flow of presentation could be kept similar as closely as possible.

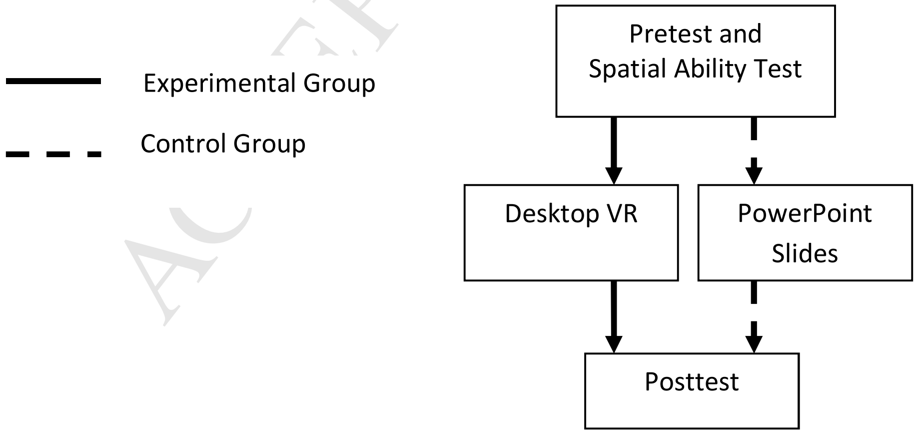


Fig. 1. Experimental process

In the ATI research, the independent variable was the learning mode which consisted of the desktop VR-based learning environment (VR mode) and the conventional classroom learning method (non-VR mode). The dependent variable was the performance achievement as measured by the posttest, while the moderator variable was the learners' spatial ability. A 2 x 2 quasi-experimental factorial design was used in which learning modes were crossed with the spatial abilities of the learners.

5.2. Population and sample

The population was senior high school science students, aged between 15 and 17 years old of any co-education high schools in a city of East Malaysia. These students were chosen because they were within the targeted populations as they have started to learn biology in senior high. Based on the simple random sampling technique, four schools were selected randomly. For each selected school, two to four intact classes were randomly chosen and assigned to either experimental or control groups.

5.3. Software

A desktop virtual reality program, V-Frog^{TM}, was used to provide the virtual learning environment to students. This software was developed and supplied by Tactus Technologies, Inc., New York. Students can cut, pull, probe, and examine a virtual specimen using the dissection tools such as scalpel and tweezer. They can use the viewpoint manipulation tools to rotate, slide and zoom the specimen. The actions are repeatable, and each dissection reflects the work of each individual student. Besides, students can get information about a part of specimen with the query tool; activate and bring parts of the specimen to life with the magic wand tool; examine an orifice in the specimen with the probe tool, and conduct a virtual endoscopy with the endoscoping tool. The existence of lab report icon on the screen indicates to students that information on the current screen can assist them in completing their lab report. Screenshots from V-Frog^{TM} are shown in Fig.2 and Fig. 3.

Fig.2. Pulling back the membrane with a tweezer to expose the internal organs (Courtesy of Tactus Technoloies)

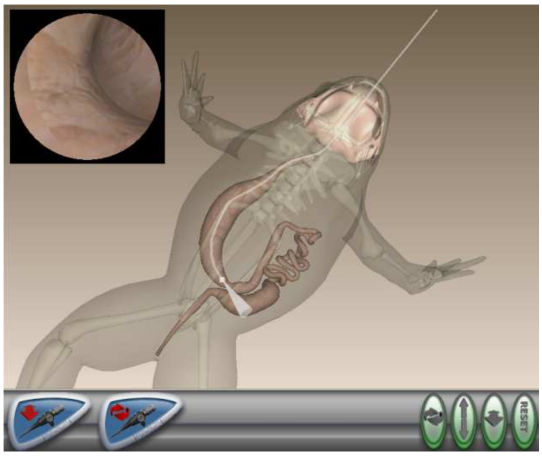


Fig. 3. Using endoscoping tool to explore the alimentary canal (Courtesy of Tactus Technology)

5.4. Instruments

5.4.1. Pretest and posttest

The content of the pretest and the posttest was similar, but the order of questions was different to avoid the set response effect. The test questions were set based on the frog anatomy modules covered in this study. The questions include sentence completion with the correct word(s); organ labeling and drawing; and multiple-choice questions. Content validity of these tests was determined by expert judgment. Three subject matter experts were requested to review the test questions and make a judgment about how well these items represent the intended content area. A pilot study was carried out in one co-education high school from the same city with forty seven randomly selected science students to obtain information that was useful to improve these tests. These included the item difficulty index, item discrimination index, and internal consistency measure. Item difficulty index is the proportion of students who answered an item correctly whereas item discrimination index measures how adequately an item discriminates between high scorer and low scorer on an entire test (Cohen, Swerdlik, & Philips, 1996). Six items were deleted in which five items were deleted due to poor discrimination and one was deleted as it had a low corrected item-total correlation (r = 0.010). As a result, the final version of the pretest and posttest contains 32 items with an alpha coefficient of 0.846. The item difficulty index was ranging from 0.27 – 0.85 which was of moderate difficulty (Hopkins, 1988).

5.4.2. Spatial ability test

Spatial ability test from Barrett & Williams (2003) was used to test the spatial ability of the participants. The Cronbach’s alpha for the spatial ability test was 0.76. It consists of 75 patterns that could be folded or formed into figures. This test explores how easily the participants can ‘see’ and manipulate shapes and figures in space (Barrett & Williams, 2003). In other words, this test evaluates a component of spatial ability called spatial visualization, which is the spatial skill needed to see and manipulate the objects in the VR-based learning environment.

First, students read instructions and viewed sample problems similar to those tested, and then they had 10 minutes to complete the test. In this study, median split was used to classified participant as having high spatial visualization ability or low spatial visualization ability. This provides a rough way to categorize learners with different spatial abilities (Mayer, 2001).

5.5. Data collection procedures

Two weeks before the treatment, respondents from both experimental and control groups were given a pretest and spatial ability test. Three modules of frog anatomy were selected for this study: Internal Anatomy, Digestive System and Circulatory System. Right before the treatment, respondents from the experimental group were given training on how to use the V-Frog^{TM} software program. During the treatment, each respondent in the experimental group was assigned to an individual computer for a self-directed lesson with V-Frog^{TM} whereas the lesson in the control group was conducted by the biology teacher with PowerPoint slides. Immediately after the treatment, which took about 1.5 hours, the respondents answered the posttest. A two-week gap between the pretest and the posttest was for the purpose of reducing the pretest sensitization threat.

5.6. Data Analysis

Frequency and proportion were used for descriptive statistics. Independent-samples t-test was conducted to determine the difference in the performance achievement which was measured by the posttest scores. Two-way analysis of variance (ANOVA) was performed to determine if any interaction existed between the learning mode and the spatial ability, related to the performance achievement which was measured by the posttest scores.

6. Results

6.1. Distribution of learners

A total of 431 students participated in this study. However, 61 participants did not fully complete all instruments. Thus, only 370 participants were taken into consideration in the analysis. Of 370 participants, 210 were in the VR mode whereas 160 were in the Non VR mode. The sample was 42% (156) and 58% (214) in males and females, respectively (Table 1). The mean age of the participants was 15.68 years old. Of all the participants in the VR mode, almost half of them, 46% have no knowledge about VR, 41% have some knowledge, only 4% know a lot more about VR and 9% have some experiences in using VR (Table 2). It is noted that the missing values was 1% because two participants did not answer this question.

Table 1
Cross tabulation of learning mode and gender

			Gender		
			Male	**Female**	**Total**
Learning Mode	VR	Count	88	122	210
		% of Total	24%	33.0	57%
	Non VR	Count	68	92	160
		% of Total	18%	25%	43%
	Total	Count	156	214	370
		% of Total	42%	58%	100%

Table 2
VR knowledge of students in the VR learning mode

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Know Nothing	96	46	46	46
	Some Knowledge	85	41	41	87
	Lots of Knowledge	8	4	4	91
	Some Experience	19	9	9	100
	Total	208	99	100	
Missing	System	2	1		
Total		210	100		

6.2. Homogeneity of Pretest

Independent-samples t-test reported no statistically significant difference in the pretest scores for the two learning modes, t(368) = 0.330, p = 0.741 (Table 3), which inferred that all learners were homogeneous in the existing knowledge of the subject matter. Statistical tests were conducted at the alpha = 0.05 significance level.

6.3. Testing of Hypotheses

Before the independent-samples t-test and two-way ANOVA were conducted, the assumptions for these tests were performed, and these tests were found to be appropriate for employment. Independent-samples t-test was conducted at the alpha = 0.05 significance level. Levene's test was used to test the homogeneity of variances. There was a significant difference in the variance of the posttest for both learning modes. Therefore, the t-value under "Equal variances not assumed" was reported. Since the Levene's test of equality of error variances shown a significant result, thus a more stringent significance level of 0.01 was used for two-way ANOVA to explore the effects of the learning mode and spatial ability on the performance achievement (Pallant, 2007).

6.3.1. Testing of H1

The statistical results rejected the null hypothesis ($p < 0.05$). Table 3 indicates that there was a significant difference in the posttest score for VR mode (M = 65.51, SD = 15.68) and Non VR group [M = 60.56, SD = 20.88; t(284.863) = 2.506, p = 0.013]. Students in the VR mode scored better in the posttest than students in the Non VR mode. Nevertheless, the magnitude of the differences in the means was small (eta squared = 0.02). This interpretation was based on Cohen's (1988) criterion: 0.01 = small effect, 0.06 = moderate effect, and 0.14 = large effect.

Table 3
Mean scores, standard deviation (SD) and t-test of pretest and posttest

Variable	VR Mode Mean (SD)	Non VR Mean (SD)	t	df	p-value	η^2
Pretest	43.14 (19.98)	42.46 (18.82)	0.330	368	0.741	-
Posttest	65.51 (15.68)	60.56 (20.88)	2.506	284.863	0.013	0.02

6.3.2. Testing of H2

As presented in Table 4, the statistical results rejected the null hypothesis (p < 0.01). The interaction effect was statistically significant [F(1, 366) = 10.75, p = 0.001]. Fig.4 reveals that the effect of learning mode was much greater for the low spatial ability group than it was for the high spatial ability group. An analysis of simple effects with independent-samples t-test as shown in Table 5 revealed the learning mode effect was not significant for high spatial group (p > 0.05), but it was significant for low spatial ability group (p < 0.05). Therefore, there was no evidence that performance achievement for high spatial ability learners in the VR mode [M = 70.35, SD = 16.05] differ from the Non VR mode [M = 71.84, SD = 14.83, t(177) = 0.629, p = 0.530]. However, there was evidence that performance achievement of low spatial ability learners in the VR mode differed from the Non VR mode. The low spatial ability learners in the VR mode [M = 60.67, SD = 13.74] scored higher in the posttest than the low spatial ability learners in the Non VR mode [M = 50.86, SD = 20.52, t(143.121) = 3.790, p = 0.0005]. The effect size was 0.06 which was of moderate effect. This means that 6% of the variance in posttest was explained by learning mode.

Table 4
Two-way ANOVA of posttest by learning mode and spatial ability

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	24650.598[a]	3	8216.866	30.600	.000	.201
Intercept	1456752.504	1	1456752.504	5425.043	.000	.937
Learning Mode (LM)	1566.787	1	1566.787	5.835	.016	.016
Spatial Ability (SA)	21277.333	1	21277.333	79.238	.000	.178
LM* SA	2885.366	1	2885.366	10.745	.001	.029
Error	98279.675	366	268.524			
Total	1608773.000	370				
Corrected Total	122930.273	369				

a. R Squared = .201 (Adjusted R Squared = .194)

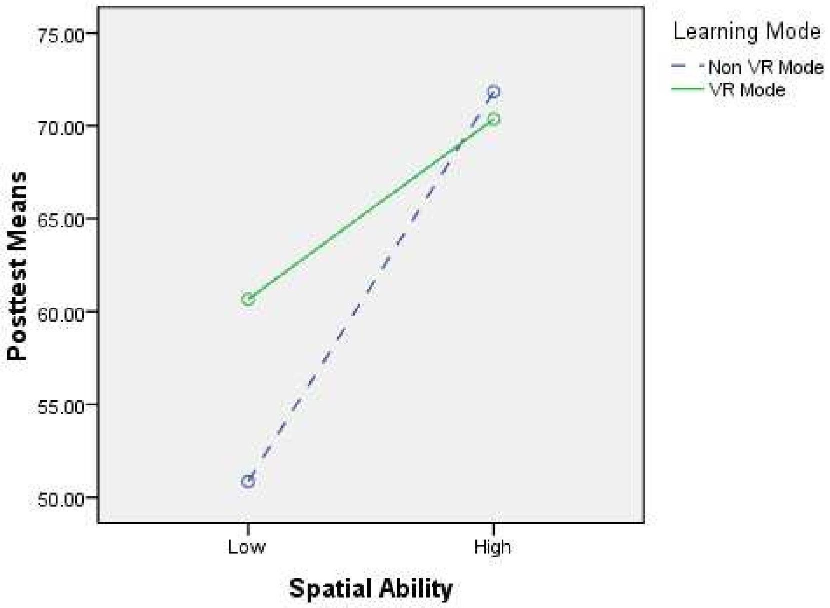


Fig. 4. Plot of interaction between learning mode and spatial ability, related to performance achievement

Table 5
T-test of posttest by learning mode for high and low spatial ability learners

Spatial Ability	Variable	VR Mode Mean (SD)	Non VR Mean (SD)	t	df	p-value	η^2
High	Posttest	70.35 (16.05)	71.84 (14.83)	0.629	177	0.530	0.002
Low	Posttest	60.67 (13.74)	50.86 (20.52)	3.790	143.121	0.0005	0.06

7. Discussion

The finding indicated better learning outcome for students in the VR-based learning environment. Based on the cognitive load theory, it is believed that students in the VR group performed better because the available memory resources were not exceeded by the total cognitive load imposed. VR instructional intervention might have helped to reduce extraneous cognitive load and at the same time increase germane cognitive load. The ability to control the learning activities while interacting with the dynamic visualizations allows learners to adapt the instructional material to their cognitive system to decrease extraneous load; and engages learners in active processing of instructional material to increase germane cognitive load (Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Khalil et al., 2005).

Moreover, spatial contiguity principle was adopted in the VR-based learning environment because the texts and graphics were placed close to each other on each screen. This appears to help decrease extraneous cognitive load as well (Ayres & Sweller, 2005; Kalyuga, 2011; Mayer, 2001). In addition, the virtual environment in the VR learning mode provides a more automated spatial encoding, and therefore does not require specific spatial processing schema to mentally transform 2-D objects into 3-D objects. Such support could help to reduce the extraneous cognitive load (Chen, 2006).

On the other hand, in the PowerPoint slides for the Non VR learning mode, due to a limited space on each page, different sources of information were separated in time. That is, texts were on one page that was presented after or before the graphic illustrations which were placed on another page. This separate design has caused the learners less likely to be able to hold both words and pictures in working memory at the same time (Mayer, 2001). The process of information integration may burden the working memory (Kalyuga, Ayres, Chandler, & Sweller, 2003). These probably explain the significant positive effects of the VR learning mode as compared with the Non VR learning mode.

However, it is noted that small effect size was found for group differences in students' performance achievement. This indicates that the result should be interpreted more cautiously in a practical sense and further replication studies should be conducted.

The ATI research has shown a significant interaction between the learning mode and spatial ability on performance achievement. Thus, performance achievement varies as a function of spatial ability and learning mode. Low spatial ability learners' performance, compared with high spatial ability learners, appeared to be more positively affected by the learning mode. On the contrary, Chen (2006) who used VR for driving instruction to high school students found no interaction effect between the learners' spatial abilities and the learning modes (Guided VR mode, Non-Guided VR mode and Non VR mode). The learners benefited most from the Guided VR mode, irrespective of their spatial abilities in Chen's study.

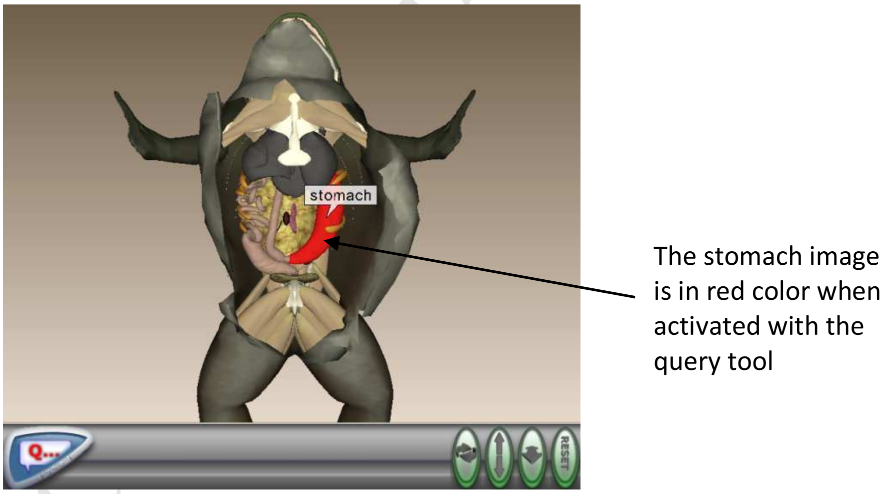
The findings of the ATI research could be explained by cognitive load theory. The major factor that contributes to cognitive load is the number of elements that a learner needs to attend to. The capability of working memory is limited to deal with two to four elements at any given time in the sense of combining, contrasting, or manipulating elements (Sweller, 2005; van Merrienboer & Sweller, 2010). The number of elements of information presented in the instructional material for the Non VR learning mode could have imposed extraneous load for the low spatial ability learners. They need to mentally transform the 2-D objects into 3-D objects while at the same time to organize, compare and contrast different organs in the digestive tract and different parts of veins and arteries in the circulatory systems. Thus, their learning was hindered because the total cognitive load was not within the confines of working memory.

The superior test performance of the low spatial ability learners in the VR learning mode compared to the Non VR learning mode indicated that the interactive virtual learning environment, animated pictures and on screen text did not provide redundant information and did not impose extraneous cognitive load. The VR-based learning has in fact managed to reduce the extraneous load for the low spatial ability learners, thus enable more working memory to be used for the processing and encoding to-be-learned information into the long-term memory. In other words, germane cognitive load occurred because free working memory resources were actively devoted to learning activities. The following reasons could further explain why low spatial ability learners performed much better in the VR learning mode.

Learner control allows low spatial ability learners to control over their interactivity with the instructional material; actively participate to search for some items in space; and closely monitor the information given to follow the lesson. These learning activities may increase the germane cognitive load to assist in schema acquisition when learning with VR. According to Hasler, Kerten and Sweller (2007, p. 725), the deeper cognitive processing of the instructional information in terms of a higher germane cognitive load is likely to result in better learning performance. In their study, learner-paced groups (segments and stop-play) have a higher test performance with relatively lower cognitive load than the system-paced group (continuous and narration-only).

In addition, split attention effect was reduced for low spatial ability learners in the VR learning mode. Split attention occurs when students need to attend to more than one source of information such as both the graphics and the text if the associated text is placed above, below or the side of the graphic. The instructional material can only be understood after the multiple sources of information is mentally integrated by the learners (Kalyuga, 2011).

A lot of information needs to be labeled in anatomical visuals. The instructional material in the VR learning mode has physically integrated the multiple resources of information and has also used visual cues such as color cueing to direct the learners' attention to the relevant parts of the diagram. For instance, in V-Frog^{TM}, the anatomical image is highlighted with a different color when it is activated with the query tool and at the same time the labeling of the image is given which is embedded onto the image as shown in Fig. 5. Hence, instructional guidance can act as a substitute to the missing schemas and help to construct schemas and automation for low spatial ability learners.



The stomach image is in red color when activated with the query tool

Fig. 5. The organ is highlighted in red and the labeling is provided when it is activated with the query tool (Courtesy of Tactus Technology)

Based on the cognitive load theory, the fact that there was no difference in performance for the high spatial ability learners in both learning modes suggested that the cognitive load imposed by both learning modes fell to a level that was within the bounds of mental resources, hence learning was not impeded. High spatial ability learners have a more expansive set of schemes on spatial intelligence and bring their activated schemas to the process of constructing mental representations of a task or situation (Cooper, 1998; Kalyuga et al., 2003). As mentioned by Cooper (1998), when learners hold high levels of expertise in the content area then their working memory may attend to elements with large complex knowledge networks. Thus, their working memory needs to attend only a few elements in order to hold all of the to-be-learned information in long-term memory. Moreover, experts can categorize multiple elements or related information as single higher level of element which requires considerably less working memory capacity for processing (Kalyuga et al., 2003). Consequently, ample cognitive resources are available for the process of learning. High spatial ability has compensated for a non interactive 3-D virtual learning environment in the Non VR mode. Hence, instructional design manipulation for this group of learners will be ineffective because their working memory capacity is not being exceeded (Cooper, 1998).

The finding that low spatial ability learners performed better in the VR mode was in line with the study of Höffler and Leutner (2011). They found that students with low spatial ability performed better with animation instructions than with static pictures which was supported by the ability-as-compensator hypothesis (Mayer, 2001). Study of Merchant et al. (2013) also identified that low spatial ability learners achieved better performance in 3-D virtual learning compared to 2-D images. The ability-as-compensator hypothesis posits that constructing mental animations from non-dynamic materials need spatial ability (Höffler & Leutner, 2011). High spatial ability learners could use their ability to compensate in an environment without explicit presentation of 3-D representation and dynamic visualization, but low spatial ability learners could not. Thus, the ability-as-compensator hypothesized that low spatial ability learners should gain particular benefit from the interactive 3-D virtual learning environment as they have difficulty to mentally construct their own visualization. The explicit presentation of 3-D representations and dynamic visualizations may keep the need for using spatial processing schema to very minimum, thus reduces the extraneous cognitive load and fosters learning.

8. Limitation and future investigation

The findings are primarily dependent on the context of biology learning. Different learning content may arouse different results. Thus, replication of the study in different learning context is recommended for future research to determine if the results can be observed in samples for other learning programs with different content and over a period of time.

Novelty effect cannot be ruled out. Almost half of the samples have never heard of VR and used VR before. Thus, this new technology may create a sense of new excitement and motivation that could influence the students' performance. Studying students in the desktop VR-based learning environment over a number of terms or semester might diminish this effect.

9. Conclusions

The significant positive effect of the desktop VR-based learning environment on the performance outcome has provided empirical evidence of the potential of desktop VR technology to support and enhance learning in biology education. This finding implies that student-centred approach in the VR learning mode is superior to the teacher-centred approach in the Non VR learning mode. The interaction study has shown that desktop VR-based learning environment benefits more to the low spatial ability learners than to the high spatial ability learners. Thus, this could help educators to facilitate individualized learning. To conclude, the findings of this research have contributed to our understanding of the potential of desktop VR in education and the learning outcome of a desktop VR-based learning environment with regard to spatial ability.

References

Allen, K., Austin, T., Beach, R., Bergstrom, A., Exon, S., Fabri, M., et al. (2002). Creating and using virtual reality: A guide for the arts and humanities. Retrieved 16 March 2008, from http://vads.ahds.ac.uk/guides/vr_guide/sect11.html

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*(1), 74-85.

Aoki, H., Oman, C. M., Buckland, D. A., & Natapoff, A. (2008). Desktop-VR system for preflight 3D navigation training. *Acta Astronautica, 63*, 841-847.

Ausburn, L. J., & Ausburn, F. B. (2004). Desktop virtual reality: A powerful new technology for teaching and research in industrial teacher education. *Journal of Industrial Teacher Education, 41*(4), 33-58.

Ayres, P., & Sweller, J. (2005). The split-attention principle. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 135-146). New York: Cambridge University Press.

Barrett, J., & Williams, G. (2003). *Test your own aptitude*. London, UK.: Kogan Page.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualizations. *Learning and Instruction, 14*(3), 325-341.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53-61.

Chen , C. J. (2006). Are spatial visualization abilities relevant to virtual reality? *e-Journal of Instructional Science and Technology, 9*(2), 1-16.

Chen, C. J., Toh, S. C., & Wan, M. F. (2004). The theoretical framework for designing desktop virtual reality-based learning environments. *Journal of Interactive Learning Research, 15*(2), 147-167.

Chen, C. J., Toh, S. C., & Wan, M. F. (2005). Are learning styles relevant to virtual reality? *Journal of Research on Technology in Education, 38*(2), 123-141.

Cheng, Y., & Wang, S.-H. (2011). Applying a 3D virtual learning environment to facilitate student's application ability – The case of marketing. *Computers in Human Behavior, 27*(1), 576-584.

Chuah, K. M., Chen, C. J., & Teh, C. S. (2008). Incorporating kansei engineering in instructional design: Designing virtual reality based learning environments from a novel perspective. *Themes in Science and Technology Education, 1*(1), 37-48.

Cohen, C. A., & Hegarty, M. (2014). Visualizing cross sections: Training spatial thinking using interactive animations and virtual objects. *Learning and Individual Differences, 33*, 63-71.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, R. J., Swerdlik, M. E., & Philips, S. M. (1996). *Psychological testing and assessment: An introduction to tests and measurement*. California: Mayfield Publishing Company.

Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW. Retrieved Sept 10, 2009, from http://education.arts.unsw.edu.au/staff/sweller/clt/

Dalgarno, B., Hedberg, J., & Harper, B. (2002). *The contribution of 3D environments to conceptual understanding.* Paper presented at the ASCILITE 2002, Auckland, New Zealand.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38*(2), 105-134.

De Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305-308.

Dünser, A., Steinbügl, K., Kaufmann, H., & Glück, J. (2006). *Virtual and augmented reality as spatial ability training.* Paper presented at the Proceedings of the seventh ACM SIGCHI New Zealand Chapter's international conference on computer-human interaction.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 117-133). New York: Cambridge University Press.

Gazit, E., Yair, Y., & Chen, D. (2006). The gain and pain in taking the pilot seat: Learning dynamics in a non immersive virtual solar system. *Virtual Reality, 10*(3-4), 271-282.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*, 713-729.

Hauptman, H. (2010). Enhancement of spatial thinking with Virtual Spaces 1.0. *Computers & Education, 54*(1), 123-135.

Hauptman, H., & Cohen, A. (2011). The synergetic effect of learning styles on the interaction between virtual environments and the enhancement of spatial thinking. *Computers & Education, 57*(3), 2106-2117.

Höffler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations – Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior, 27*(1), 209-216.

Hopkins, K. D. (1988). *Educational and psychological measurement and evaluation* (8th ed.). USA: Allyn & Bacon.

Huang, H.-M., Rauch, U., & Liaw, S.-S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education, 55*(3), 1171-1182.

Huk, T. (2006). Who benefits from learning with 3D models? The case of spatial ability. *Journal of Computer Assisted Learning, 22*, 392-404.

Inoue, Y. (2007). Concepts, applications, and research of virtual reality learning environments. *International Journal of Social Sciences, 2*(1), 1-7.

Kalyuga, S. (2011). Cognitive load theoy: How many types of load does it really need? *Educational Psychology Review, 23*, 1-19.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38* (1), 23-31.

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55*(2), 427-443.

Khalil, M. K., Paas, F., Johnson, T. E., & Payer, A. (2005). Design of interactive and dynamic anatomical visualizations: The implication of cognitive load theory. *The Anatomical Record (Part B: New Anat.), 286B*, 15-20.

Lajoie, S. P. (2008). Individual differences in spatial ability: Developing technologies to increase strategy awareness and skills. *Educational Psychologist, 38*(2), 115-125.

Lee, E. A.-L., & Wong, K. W. (2008). A review of using virtual reality for learning. *Transactions on Edutainment, 1*, 231-241.

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2009). Learning effectiveness in a desktop virtual reality-based learning environment. In S. C. Kong, H. Ogata, H. C. Arnseth, C. K. K. Chan, T. Hiroshima, F. Klett, J. H. M. Lee, C. C. Liu, C. K. Looi, M. Milrad, A. Mitrovic, K. Nakabayashi, S. L. Wong & S. J. H. Yang (Eds.), *Proceedings of the17th International Conference on Computers in Education [CDROM]*. Hong Kong: Asia-Pacific Society for Computers in Education.

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010a). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*(4), 1424-1442.

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010b). Learning with virtual reality: Its effects on students with different learning styles. *Transactions on Edutainment, IV*, 79-90.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merrienboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058-1072.

Mayer, R. E. (2001). *Multimedia learning*: Cambridge University Press, Cambridge, UK.

McGee, M. G. (1979). *Human spatial abilities, sources of sex differences*: Praeger, N. Y. .

Meijer, F., & van den Broek, E. L. (2010). Representing 3D virtual objects: Interaction between visuo-spatial ability and type of exploration. *Vision Research, 50*, 630-635.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29-40.

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Cifuentes, L., Kwok, O., & Davis, T. J. (2013). Exploring 3-D virtual reality technology for spatial ability and chemistry achievement. *Journal of Computer Assisted Learning, 29*(6), 579-590.

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O.-m., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education, 59*(2), 551-568.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education, 56*(3), 769-780.

Pallant, J. (2007). *SPSS survival manual* (3rd ed.). Sydney, Australia: Allen & Unwin.

Petersson, H., Sinkvist, D., Wang, C., & Smedby, Ö. (2009). Web-based interactive 3D visualization as a tool for improved anatomy learning. *Anatomical Sciences Education, 2*, 61-68.

Plass, J. L., Kalyuga, S., & Leutner, D. (2010). Individual differences and cognitive load theory. In J. L. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 65-87). New York: Cambridge University Press.

Strangman, N., & Hall, T. (2003). Virtual reality/simulations. Wakefield, MA: National Center on Accessing the General Curriculum. Retrieved May 9, 2006, from http://www.cast.org/publications/ncac/ncac_vr.html

Stull, A. T. (2009). *Anatomy learning in VR: A cognitive investigation.* (Ph.D.), University of California, Santa Barbara.

Stull, A. T., Barrett, T., & Hegarty, M. (2013). Usability of concrete and virtual models in chemistry instruction. *Computers in Human Behavior, 29*(6), 2546-2556.

Sweller, J. (2005). Implications of cognitive load theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19-30). New York: Cambridge University Press.

Sweller, J. (2010a). Cognitive load theory: Recent theoretical advance. In J. L. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 29-47). New York: Cambridge University Press.

Sweller, J. (2010b). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123-138.

Tactus Techologies. (2007). V-Frog: Virtual frog dissection. Retrieved May 22, 2008, from http://www.tactustech.com/products/vfrog/documents/V-Frog-White-Paper.pdf

Urhahne, D., Nick, S., & Schanze, S. (2009). The effect of three-dimensional simulations on the understanding of chemical structures and their properties. *Research in Science Education, 39*, 495-513.

van Merrienboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education, 44*(1), 85-93.

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177.

Wang, H.-C., Chang, C.-Y., & Li, T.-Y. (2007). The comparative efficacy of 2D-versus 3D-based media design for influencing spatial visualization skills. *Computers in Human Behavior, 23*, 1943-1957.

Wu, Y., Yuan, Z., Zhou, D., & Cai, Y. (2013). A Mobile Chinese Calligraphic Training System Using Virtual Reality Technology. *AASRI Procedia, 5*, 200-208.

Highlights

- Desktop VR-based learning environment shows positive effect on students' performance.
- Performance achievement varies as a function of spatial ability and learning mode.
- Low spatial ability learners benefit more in VR learning mode.