

ORIGINAL RESEARCH

Does ‘WOW’ translate to an ‘A’? Exploring the effects of virtual reality assisted multimodal text on Chinese Grade 8 EFL learners’ reading comprehension

Yudan Su

Shenzhen Middle School, Shenzhen, Guangdong Province, China

Received: November 7, 2021 / Accepted: December 1, 2021 / Published Online: December 2, 2021

© Pioneer Publications LTD 2021

Abstract

Purpose In recent years, the incorporation of multimedia into linguistic input has opened a new horizon in the field of second language acquisition (SLA). In the reading aspect, the advent of virtual reality (VR) technology extends the landscape of reading repertoire by engaging learners with auditory, visual and tactile multimodal input. This study aimed to examine the pedagogical potential of VR technology in enhancing learners’ reading comprehension.

Methods Three classes including 131 Chinese 8th grade EFL students participated in this study. This study adopted mixed methods methodology and triangulated pre-post-retention tests, questionnaires, learning journals and interview data to compare three modes of text input on learners’ reading performance and cognitive processing. **Results** The results indicated that VR-assisted multimodal input significantly improved learners’ macrostructural comprehension in the short term, whereas there was no significant difference of retention performance. The findings revealed that reading multimodal text did not exceed learners’ memory capacity or impose extraneous cognitive load. Participants mainly reported favorably on the efficacy of multimodal input in assisting their reading. **Conclusion** This study was the first attempt to integrate VR technology with input presentation and cognitive processing and offered a new line of theorization of VR-assisted multimodal learning in the cognitive field of SLA.

Keywords multimodal input; virtual reality; reading comprehension; cognitive load

1. Introduction

In the field of second language acquisition (SLA), the important role of input has been addressed by many researchers (Ellis, 1994; Gass, 1997; Krashen, 1985; Long, 1996). In the cognitive account of SLA, the exposure to input has been regarded as a necessary condition for second language (L2) development to occur (Fotos, 2000; Gass, 1997). Recently, the modality of input has become a focus of inquiry in SLA due to the development of information and communication technology (ICT), which has transformed the way information is recorded, represented, managed and processed.

In this study, I used the term “multimodal text” (Walsh, 2007, p. 26) to encompass a broad concept of engaging in, interacting with, and reflecting upon the text presented by different multimedia. This study was situated in the cognitive field of SLA to test the efficacy of multimodal input on Chinese EFL beginners’ reading comprehension by comparing screen-based multimodal

text and print-based monomodal text, specifically how learners decode word meanings and memorize details at the microstructural level and construct coherent mental representations at the macrostructural level.

The role of multimodality in SLA has been emphasized by advocates of multimedia learning (Rost, 2002), since multimodality provides learners with multisensory information in diverse semiotic codes (Legros and Crinon, 2002). In the past decade, two-dimensional (2D) visuals such as pictures and videos have been explored extensively in providing multimodal input (Lorenz, 2009; Lan and Sie, 2010). The advent of sophisticated virtual reality (VR) technology extends the scope of 2D multimodal input to a three-dimensional (3D) level and engages learners with auditory, visual and tactile multimodal input. Understanding how the affordances provided by VR technology affect language learning and how individuals learn with the assistance of VR is not well understood. This research is expected to take forward the field by providing a pedagogical rationale that understands how students interact with VR-

assisted multimodal text and thereby improving reading performance.

This study applied VR technology to examine one facet of SLA: reading, particularly the expository reading that situates readers in a content and language integrated learning (CLIL) context. For some students, reading expository text can be an arduous and occasionally frustrating experience (Lightbown & Spada, 2013). One of the underlying reasons for the current study was to examine one potential avenue that could make the reading process more enjoyable, effective and efficient.

This study looked at the role of working memory to evaluate whether multimodal input facilitates learners' reading comprehension by exerting modality effect within limited capacity or hinders learners' comprehension by imposing redundancy effect exceeding memory capacity. Thus, the cognitive load approach was also adopted (Paas et al., 2003) to compare learners' cognitive load in three input contexts and explore their perceptions of multimodal input to capture a full picture of how multimodal text affects EFL learners' reading development. To my knowledge, there has been little attempt to compare the effects of multimodal input and monomodal input in the reading aspect of SLA, and this study addressed this gap by making comparisons in three aspects: (1) reading comprehension at macro-structural and microstructural levels in different input conditions, (2) cognitive load imposed by different presentation modes, and (3) learners' perceptions of the multimodal text in the reading treatment. Overall, the study aimed to test the efficacy of multimedia, especially VR technology, in providing multimodal input and enhancing Chinese 8th grade EFL students reading comprehension, and by doing so, extend existing theories of multimedia learning and offer valuable insight of multimodality in the scholarship of SLA.

2. Theoretical background

In general, there are two major strands of theories that link SLA with multimedia learning. The first strand is based on the established Krashen's input hypothesis in SLA while the second perspective draws on a cognitive framework of multimedia learning. In the field of SLA, a substantial amount of empirical studies utilized multimedia to optimize linguistic input, while the cognitive theory has triggered more inquiry into learners' inner mechanism. In this study, the two theoretical perspectives were brought together to conceptualize a framework for multimodal reading assisted by VR technology.

2.1. Input hypothesis in SLA

Multimedia learning is related to a number of SLA theories, one of which is Krashen's input hypothesis, indicating that multimedia tools can be incorporated in the process of L2 learning through the combination of different modes of input (Wang, 2012).

According to Krashen (1981, p. 104), "the optimal

input is slightly above the present level of learners' competence as an 'i + 1' model." However, in Krashen's account, the scope of the 'i + 1' input is still not clear, and several questions can be raised, such as what degree of increase in difficulty is suitable, and whether the 'i + 1' model is applicable to all EFL learners with different levels of English proficiency. In the context of multimedia learning, input is perceived through both auditory and visual channels, and therefore both words and images are selected to create mental models of language and content. Pictures are connected to build a pictorial model, and words are connected to build a verbal model. To avoid oversimplification, this study moved one step further by utilizing VR technology to provide tactile input as 'i + 1' on the basis of auditory and visual input and examined whether it could improve a group of 8th grade Chinese EFL learners' reading comprehension within limited memory capacity or exceed their capacities and impede learning. Moreover, the acquisition process is not clearly illustrated in this hypothesis. Krashen (1982, p.21) simply claimed that "a necessary (but not sufficient) condition to move from stage 'i' to stage 'i + 1' is that the acquirer understands the 'i + 1' input." It is arguable whether understanding input alone is enough for acquisition. Hence, the cognitive account in the field of multimedia learning can make up for the deficiency by providing a detailed explanation of cognitive process.

2.2. Cognitive theory of multimedia learning

Mayer's (2002) cognitive theory of multimedia learning (see Figure 1) is underpinned by three assumptions from cognitive science: (1) the dual channels assumption – there are two separate visual and auditory channels for processing different types of information; (2) the limited capacity assumption – there is limited capacity in each channel to process information; and (3) the active processing assumption – learning is an active process that filters, selects and organizes new information and integrates it with prior knowledge. The memory system consists of three storage structures: sensory memory, working memory and long-term memory. Sensory memory acts as a buffer for stimulus received from different modes of input, working memory is short-term memory for temporary retrieval of the processed information, and long-term memory keeps large amount of information over a long period of time. According to Mayer (2009), working memory plays a key role in multimedia learning. Likewise, Sweller et al. (2011) argued that the cognitive load imposed on working memory should be taken into consideration when designing multimedia learning environment since the selection, organization and integration of information occur in working memory. Therefore, it is significant to examine the cognitive load imposed by different modes of input and evaluate how it affect learners' reading comprehension. There are three types of cognitive load distinguished in the literature: intrinsic, extraneous, and germane cognitive load (Brunken et al., 2003; Sweller et al., 1998). Intrinsic load is

attributed to the inherent difficulty level of learning material without being affected by instructional design, that is, the consistent level of difficulty of expository text in this study; extraneous load refers to the mental load caused by the presentation format and instructional design, and this is the key aspect this study looked into since it concerns with input modality; germane load results from appropriate instructional design and helps learners construct and process schemas of input. This study focused on the modality of delivering the information and examined whether multimodal input would incur extraneous load and exert the redundancy effect that affects students' performance negatively or increase the germane load and exert the modality effect that enhances their learning outcome.

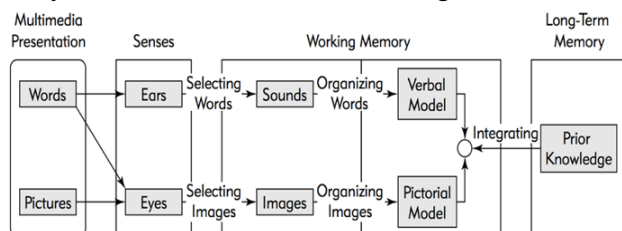


Figure 1. Cognitive theory of multimedia learning (Mayer, 2002)

Critics of this theory often question whether cognition is mediated by something other than words and images (Reed, 2010), and this study answered the question by incorporating tactile input into the framework. With the advent of technologies such as 3D modelling and VR platforms, the possibilities of multimedia learning expand exponentially. Moreno (2006) expanded Mayer's framework (2002) to include "media such as virtual reality, agent-based, and case-based learning environments" (p. 313) by adding manipulative input on the presentation end and constructing tactile sensory memory in the memory system (see Figure 2). The haptic feature of VR technology allows learners to interact with the virtual world and reinforce the information through the third sensory channel on the

basis of the dual channels. In this light, VR-assisted multimodal input can provide learners with auditory narration, visual presentation and tactile interaction and promote learners' active processing through the triple memory model in the multimodal learning context. However, this model remains vague about how different multimodal input enter in working memory and construct mental representations by selecting, connecting and organizing information. Therefore, I reconceptualized the framework for the current study by illustrating the working memory part clearly and integrate it with the input hypothesis in SLA.

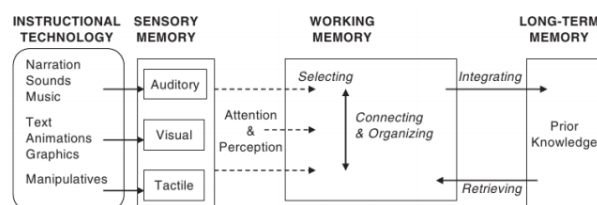


Figure 2. Cognitive theory of learning with media (Moreno, 2006)

2.3. Integrated model of cognitive theory of learning with VR

The current research combined the aforementioned theoretical perspectives and conceptualized an integrated model of cognitive theory of learning with VR. Figure 3 models the detailed learning process in the VR-assisted multimodal learning environment, which extends the breadth and depth of learners' exposure to the target text. This model also provides detailed explanations of cognitive processing in terms of mental representations and constructions. The central concept of this theory taps into the input hypothesis in SLA, the human cognitive processing system and the cognitive load principles in providing three modes of input for effective learning without exceeding working memory capacity.

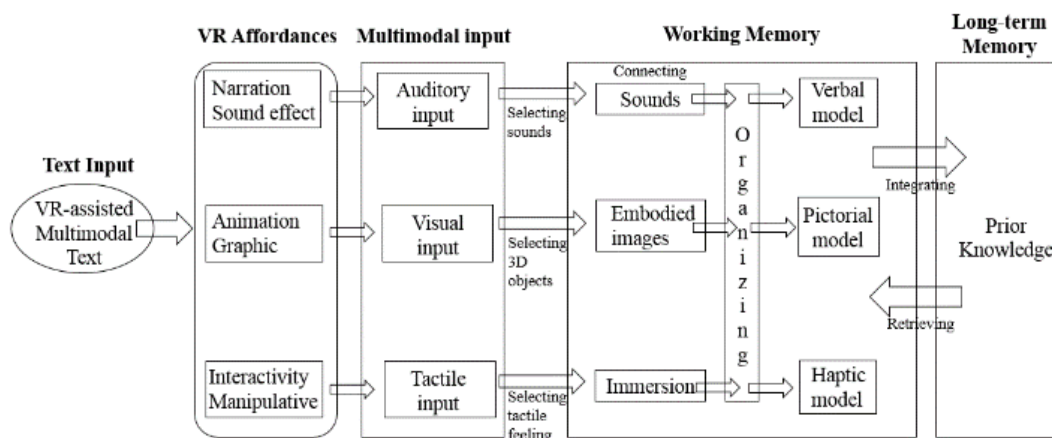


Figure 3. Integrated model of cognitive theory of learning with VR

Based on the integrated model, learners firstly pay attention to auditory, visual and tactile input attributed

to VR affordances, and then process the multimodal information actively in working memory and mentally

organize it into verbal, pictorial and haptic models respectively. Finally, the multimodal text input is integrated with existing knowledge and stored in the long-term memory. It is hypothesized that engaging in such cognitive processes in VR-assisted multimodal learning environment enables learners to construct “a coherent mental representation that integrates the textual information and relevant background knowledge” (van den Broek, 2010, p. 453) within memory capacity, and thereby leading to effective learning.

3. Literature review

3.1. Effects of multimodal input on learners’ reading performance

Some studies confirmed the modality effect of multimodal input on learner’s reading performance, and most of which focused on visual and auditory input. Different types of visual input have been proved to be beneficial to learners’ reading comprehension. Son (2003) investigated effects of three different types of text formats on learners’ comprehension, and the finding showed that computer-based hypertext format paved the way for greater comprehension than paper-based format and computer-based non-hypertext format texts. According to Pearman and Lefever-Davis (2006), CD-ROM storybooks improved school children’s overall reading comprehension because students could listen to the vivid narration of the story. In addition, some studies focused on the synergy exerted by the combination of visual and verbal input on learning outcome. Segers and Hulstijn-Hendrikse (2008) investigated the effects of dual input on EFL beginners’ cognitive processes in the multimedia learning context, and the result indicated that students who utilized oral presentation with pictures performed better than their counterparts who used written presentation with pictures. Similarly, However, few studies have confirmed the facilitative effects of VR technology in assisting L2 reading. One example is Dev, Doyle, and Valente’s study (2002) that adopted Orton-Gillingham technique to provide visual, auditory, and kinesthetic multimodal input to assist special children’s reading. The findings showed that the multimodal approach helped children improve their reading abilities out of the special level and the gains were maintained after even two years (Dev et al., 2002).

In contrast, some studies were not able to validate the facilitative effects of multimedia on learners’ reading performance. According to Rasch & Schnotz (2009), research findings were not able to show that students learned better from text and pictures than from text alone, calling the multimedia principle and the cognitive theory into question. Furthermore, Mangen, Walgermo, and Bronnack (2013) compared the effects of electronic text reading in PDF and paper text reading on tenth graders’ reading comprehension in Norway, showing that students who received paper text achieved better reading outcome than the electronic group.

The mixed experimental results regarding the effects of multimodal input on learners’ reading comprehension calls for further examination. It is noted that the majority of previous studies was limited in providing auditory and visual input, and to date there exists a paucity of studies examining the use of VR-assisted multimodal text in the context of L2 reading and no study has focused on Chinese EFL beginners’ expository reading comprehension in a CLIL context. Therefore, this study will address these research gaps by testing the efficacy of VR-assisted multimodal input on Chinese 8th grade EFL learners’ macrostructural and microstructural reading comprehension

3.2. Effects of multimodal input on learners’ cognitive load

Some studies confirmed the modality effect of multimedia in lowering learners’ cognitive load and improving their learning outcome. Lin & Yu (2012) carried out an experiment via mobile phones in Taiwan, and they divided multimodal input into text mode, text-audio mode, text-picture mode and text-audio-picture mode. The results showed that the text-audio-picture mode imposed lower mode than the others, confirming that modality effect facilitated language learning. Similarly, McClean et al. (2005) argued that animations in the lecture allowed students to process information using the two channels and reduced their cognitive load, thereby improving their retention of biological text.

Conversely, redundancy effects were also found in conditions where there were duplicated information, logically unrelated instructional material, and complex content in the multimedia-assisted learning environment (Kalyuga, Chandler, & Sweller, 2000; Sweller et al., 2011). Kalyuga et al. (1999) found that the use of simultaneous duplicated information generated additional cognitive load while information presented in only auditory format rendered performance effective. In a similar vein, Liu and Su (2011) found that simulations loaded with multimedia features increased learners’ cognitive load and learners failed to integrate information properly.

3.3. Learners’ and teachers’ perceptions towards multimodal input

Recently, the qualitative strand of research in multimedia learning has enriched the field by providing explanation or posing challenges to quantitative experimental findings. Along these lines, it is worth mentioning that some studies (Ayres, 2002; Heller, 2005; Neo, 2009; Stepp-Greany, 2002; Wiebe & Kabata, 2010) indicated that learners held positive attitudes towards the integration of multimedia with language learning. Neo (2009) investigated Malaysian students’ perceptions in a multimedia project, showing students’ positive attitudes with respect to their language learning motivation and teamwork abilities. Nair (2012) applied VR technology in an experiment and found that learners held positive attitudes towards its usefulness as a learning tool. As for teachers’ perceptions, Al-Seghayer (2016)

assessed English instructors' perceptions towards the effectiveness of electronic text on learners' L2 reading performance. Results showed that instructors held positive attitudes towards the electronic text because it improved accessibility, readers' interaction with the text and stimulated learners' interests in reading. However, few studies inquired into both learners and instructors' perceptions of multimodal input, and even few revealed negative comments of multimodal learning. Thus, this study is of conceptual value of capturing both learners and teachers' positive and negative comments towards multimodal input in the context of L2 expository reading in a comprehensive manner.

To sum up, research so far yielded conflicting findings regarding the efficacy of multimodal input on learners' reading comprehension, especially when multimedia tools were directly compared with traditional print medium. This points out the need for further investigation of how multimodal input affects learners' reading comprehension from multiple perspectives. The present study distinguished itself from previous studies in four aspects. Firstly, multiple studies have investigated different multimedia tools such as pictures, audio and video to facilitate learners' L2 acquisition, while VR technology has not been fully explored in language instruction, especially in the expository reading aspect and in Chinese educational setting. Secondly, most studies assessed learners' overall reading performance without examining different aspects. This study divided reading comprehension into microstructural and macrostructural levels and presented a detailed understanding of multimodal input's role in assisting learners' two levels of text processing. Thirdly, this study was the first attempt to examine the effects of VR-assisted multimodal input on learners' reported

levels of cognitive load including mental load and mental effort and evaluate the effectiveness of multimodal input on L2 reading from the perspective of cognitive load. Lastly, this study probed into learners' and teachers' subjective cognition through semi-structured interviews and learning journals besides objective performance to provide an extensive and intensive understanding of the efficacy of multimodal input in enhancing Chinese EFL learners' reading comprehension.

3.4. Research questions

Based on the integrated framework of cognitive theory of learning with VR, this study attempted to answer the following questions:

1. What are effects of input modalities (VR-assisted multimodal text, video-assisted multimodal text, print-based monomodal text) on Chinese EFL learners' reading performance?
2. What are effects of input modalities (VR-assisted multimodal text, video-assisted multimodal text, print-based monomodal text) on Chinese EFL learners' cognitive load?
3. What are learners and teacher's perceptions towards multimodal text in assisting reading comprehension?

4. Research Design

To address the three research questions, this study adopted a mixed methods research methodology under the guidance of pragmatist paradigm. The mixed methods research design can be briefly summarized in the Table 1 to answer the three research questions.

Table 1. The integrated mixed methods research design

	Focus		Data collection	Data analysis
RQ1: the effects of input modalities on learners' reading performance	Immediate effects	Overall comprehension	Immediate post-test	SPSS: MANOVA, ANOVA
		Macrostructural comprehension		
		Microstructural comprehension		
	Delayed effects	Overall comprehension	Delayed post-test	
		Macrostructural comprehension		
		Microstructural comprehension		
RQ2: the effects of input modalities on learners' cognitive load	Cognitive load	Overall load	Cognitive load scale	SPSS: One-way ANOVA
		Mental load		
		Mental effort		
RQ3: learners' and teacher's perceptions about the efficacy of multimodal text	Learners' perceptions		Focus group interview, learning journals	Coding: Inductive and deductive coding
	Teacher's perception		Individual interview	

4.1. Research paradigm

This study is situated in the pragmatist paradigm which "is not committed to any one system of philosophy or reality but focuses on the 'what' and 'how' of the research problem" (Creswell, 2003, p.11). Pragmatism allows independent collection of quantitative and qualitative data and integration of the two strands at the stage of interpretation and inference (Tashakkori & Teddlie, 1998). As shown in Figure 4, the study utilized

concurrent triangulation strategy by collecting quantitative and qualitative data respectively and synthesizing findings at the interpretation stage. This strategy is an optimal approach because it costs less time to collect both strands of data in comparison to the sequential method.

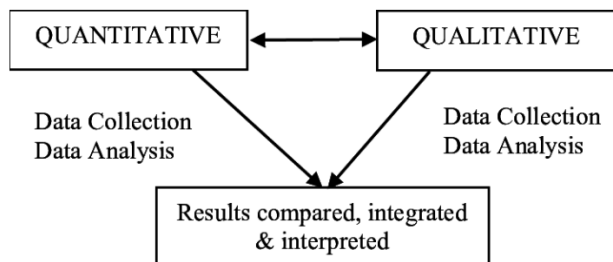


Figure 4. Visual diagram of mixed methods concurrent triangulation strategy (Atif et al., 2013)

This study adopted quasi-experimental research design to fully gauge the efficacy of VR-assisted multimodal text input on learners' reading performance. This study tackled the validity issue by selecting three classes at similar level of average academic performance and language proficiency, since the target school has streamed students into three levels (above average, average and below average) based on their academic performance. It is also noted that the period of treatment was short and students' performance could be atypical and unnatural under experimental condition. Thus, the current research also collected a qualitative strand of data out of the class setting, which was conducive to providing a comprehensive understanding of the effects of multimodal input on Chinese EFL learners' reading comprehension.

4.2. Methods

4.2.1. Research site and access

The research site was a middle school located in Jiangxi Province, China. The provincial government and the Ministry of Industry and Information Technology in Nanchang, Jiangxi Province, are trying to take the lead in building the city as a world-class VR center and provide support for schools to implement the advanced technology. The research site was the first school that has applied VR technology into secondary education and built a VR lab for instructional use. Moreover, this school has incorporated VR lessons in the curriculum to teach biology, geography and history classes in the VR lab on a weekly basis, whereas the English subject has not been incorporated in VR lessons yet.

4.2.2. Participants and sampling

The sample in this study was 8th grade Chinese L1-English L2 learners in the target school. A total of 137 students in grade 8 participated in the study while six students were excluded from the sample because they did not finish either immediate post-test or delayed post-test. As shown in Table 2, the sample was comprised of three classes (N = 131), with 42 students in experimental group A, 46 students in experimental group B, and 43 students in control group C. In the sample, males (n = 64) represented a smaller proportion than females (n = 73). Participants' ages ranged from 12 to 15, with a mean of 14 years (SD = .536). Chinese was their first language, and 96% of students have learned English for more than five years. A total of 62

(45.3%) participants found reading was the most difficult aspect of learning English, especially expository text due to technical vocabulary and unfamiliar content. In terms of multimedia, teachers often use interactive whiteboard to present power point or play videos in class, and 72.3% of participants evaluated it helpful for understanding. There were no significant differences found across the three experimental groups with respect to gender, age or time of learning English. In addition, another important participant is Ms Li, the English teacher of the three classes, who has over 12 years of teaching experience with a master's degree in English education and previous experience of incorporating multimedia tools in English teaching, and she also knows how to operate the research apparatus. Ms Li followed the whole experiment and shared her own insight in the individual interview.

Table 2. Participants' profile

Category	Group A (n = 42)		Group B (n = 46)		Group C (n = 43)	
	Freq	%	Freq	%	Freq	%
Gender						
Female	23	54.76	24	52.17	22	51.16
Male	19	45.24	22	47.83	21	48.84
Age (years)						
12	0	0	0	0	1	2.33
13	16	38.1	21	45.65	17	39.53
14	25	59.52	24	52.17	25	58.14
15	1	2.38	1	2.17	0	0
Time of learning English (years)						
3-5	10	23.81	12	26.09	19	44.19
6-8	31	73.81	31	67.39	22	51.16
9-11	1	2.38	3	6.52	2	4.65
Difficult aspect						
Listening	1	2.38	2	4.35	4	9.3
Reading	19	45.24	18	39.13	24	55.81
Writing	15	35.71	10	21.74	9	20.93
Speaking	7	16.67	16	34.78	6	13.95
Usefulness of multimedia						
Useful	35	83.33	32	69.57	28	65.12
Not useful	2	4.76	4	8.7	3	6.98
Not sure	5	11.9	10	21.74	12	27.9

This study adopted non-probability purposive sampling technique due to the low feasibility in drawing a random sample from all 8th grade EFL learners in the target school. To address the threat to the research validity, three classes at the same level of academic performance were selected and assigned into two experimental groups and one control group. Experimental group A read VR-assisted multimodal text with visual, auditory and tactile input, experimental group B read video-assisted multimodal text with visual and auditory input, while the control group C read print-based monomodal text with visual input only. Since the multimodal reading sessions were designed to be learner-centered, Ms Li only led the reading activity with

minimal involvement, thereby avoid interfering with learners' reading in the intervention.

4.2.3. Research apparatus and treatment materials

The research apparatus used in this study is zSpace all-in-one computer. zSpace is an interactive hardware and software platform where learners can listen, watch and feel the multimodal input that cannot be achieved in a conventional computer environment. As shown in Figure 5, it mainly consists of three components: an all-in-one computer with a 24-inch 3D stereoscopic display, three pairs of polarized glasses, and a laser-based interactive stylus. The screen has built-in tracking sensors to trace viewing angles of readers, and it is installed on a stand which tilts it up around 30° for learners to observe. The stylus pen possesses three buttons, one for learners to select the objects shown on the screen, and the other two for learners to zoom in and out to observe the 3D model in a full view. The three components can be activated simultaneously to situate readers in an immersive and interactive learning environment. This study used all-in-one VR computer rather than common VR headsets because it allowed students to interact with peers and teachers in the classroom setting rather than completely absorb themselves in the virtual world.



Figure 5. Major components of the zSpace platform

zSpace integrates visual, auditory and tactile elements and offers learners multimodal input. I want to use one reading topic 'water's journey' to demonstrate how zSpace was applied in practice. For experimental group A, the VR-assisted multimodal text was composed of visual input which showed 3D animation of the water cycle and content-related pictures along with words, auditory input which narrated the digital content on the screen with acoustic effects like raindrop and water boiling, and haptic input which allowed learners to feel the water flow as if learners were experiencing it out of the device. For experimental group B, the video-assisted multimodal text was made up of visual and auditory input without haptic feedback. Participants watched the video clips that showed the same textual information of the reading topic with English subtitles as digital text and the cartoon character's narration and sound effects as auditory input in a classroom

equipped with projector and computer. Students in the control group C only received the visual input of print-based text. A glossary of new words was given in all three groups.

The reading materials used in the treatment were extracted from past test papers in senior high school entrance examination. The justification for choosing these texts is twofold. Firstly, the past exam papers are widely used in 9th grade as sample test, and students in 8th grade generally do not have access to them, which ensure that participants are at the same baseline without prior knowledge of reading tasks. Although some reading texts may be demanding for 8th grade students, the Chinese annotation of new words in the text was given to lower the level of difficulty in accordance with students' level of competence. Secondly, passages used in the authoritative examination were carefully selected and reviewed by experienced English teachers and the Ministry of Education. Thus, the validity of using these tasks to assess learners' reading comprehension was promised. The selection of treatment materials also took the availability of the same content in both VR and video platforms into consideration. As a result, I prepared six expository texts in English of comparable length (200-250 words), general science topics (e.g. water's journey, butterfly's lifecycle, frog's lifecycle) and similar structure with four to five graphs, the last one being a summary of the main idea. Two of these texts were used in the pre-test, two in the immediate post-test and two in the delayed post-test.

4.2.4. Data collection

The study mainly used reading tests to obtain the objective knowledge of learners' reading performance. The reading test was formatted in multiple-choice questions and blank-filling questions to evaluate learners' reading performance in an objective manner and minimize the potential threat of subjective grading to the research validity. In each reading task, there were five questions in total, with three questions testing microstructural understanding and two examining macrostructural understanding. Microstructural understanding was captured by readers' ability to answer questions based on explicitly stated details in the text correctly (e.g. the correct explanation of butterfly's metamorphosis), whereas macrostructural understanding was assessed by questions on text summary and implications (e.g. the possible effect if water is contaminated in the transmission stage).

Students had to complete reading tests three times as pre-test, immediate post-test and delayed post-test respectively. Prior to the intervention, students completed a pre-test including two paper-based expository texts in the intervention and a total of ten questions, and the average score of each component formed the baseline data of participants' expository reading ability at macrostructural and microstructural levels. In addition, the pre-test assessed learners' prior knowledge of reading materials, and the variable of prior knowledge was controlled by removing learners who had sufficient domain-specific knowledge before the intervention. After

receiving different modes of input in the intervention, participants were required to complete the immediate reading test. The average score of two post-tests after reading sessions was regarded as the immediate post-test result. The delayed post-test was administered two weeks after the intervention and students had to finish the test without the aid of multimodal text. The delayed post-test scores reflected learners' retention of textual information, and were compared with pre-test and immediate post-test scores to examine the long-term effects of the multimodal reading treatment. All marking was completed without group membership by English teachers with no prior knowledge about the experiment to avoid any bias towards one group or the other.

This study utilized the survey instrument in two ways. Firstly, prior to the treatment, a demographic questionnaire (see Appendix 1) was administered to get a snapshot of participants' background information and allow me to have a general understanding of the sample. The participants' profile showed that there was no significant difference among three groups with respects to age, gender, time of learning English and attitudes towards multimedia learning. Secondly, participants were required to complete a cognitive load questionnaire after each reading session to report their invested mental load and effort in reading the expository text. This cognitive load scale has been widely used in literature and this study adapted the scale from the measures of Paas (1992) and Sweller, van Merriënboer, and Paas (1998) and Hwang, Yang, and Wang (2013). This questionnaire consisted of eight items in mental effort and mental load dimensions with a five-point Likert rating scale (see Appendix 2). The Cronbach's alpha was applied to ensure the satisfactory reliability of this instrument. The result of cognitive load ratings was evaluated on a group basis and compared under three input conditions to examine the interrelation between input modality and cognitive load.

This study investigated learners' perceptions towards the efficacy of multimodal input through collecting learning journals extensively. Learning journal is one common approach of collecting data in qualitative research (Janesick, 1999) and considered as an effective way to obtain learners' perceptions (Cohen, Manion, & Morrison, 2007). In this study, participants were asked to write a learning journal after each session to provide narrative accounts of their perceptions towards multimodal text as part of their learning experience. The journal topics were based on the coding scheme regarding perceptions of different modes of input and overall reading experiences with multimodal text. The purpose of the journals was to gain a contextual understanding of the participants' experiences in reading expository texts with multimodal input. Additionally, the collection of every participant's journal entry widened the amount of qualitative data besides in-depth interview, and it gave students who preferred to writing rather than talking an alternative way of sharing their thoughts and attitudes of multimodal input in the reading intervention.

This study approached teacher and learners'

perceptions by conducting focus-group interviews and individual interview. This study utilized focus group interview method to bring group of 6-8 people together to discuss a shared experience (Creswell, 2003). In the current research, the focus group interview was conducted with six students from experimental groups in a semi-structured way. Each group interview lasted for around 20 minutes and all interviews were audio-recorded and transcribed for analysis with interviewees' permissions. In addition, I had an individual interview with Ms Li after the treatment and explored how the experienced English teacher perceive the effects of multimodal text on students' L2 reading comprehension. Thus, both insider and outsider perspectives towards the efficacy were gained from conducting interviews in depth. The interview questions (see Appendix 3) have been checked by two English teachers and applied in the pilot study to ensure the reliability of the instrument. Interviews were conducted in participants' L1 Chinese according to their own preferences so that interviewees could share opinions at ease and the reliability of qualitative findings could be strengthened.

Throughout the data collection process, all participants remained anonymous. The data collected by the three instruments were triangulated to test the efficacy of multimodal input on Chinese 8th grade EFL learners' reading comprehension.

4.2.5. Research procedure

The entire research procedure can be largely divided into three stages: pre-intervention, reading intervention and post-intervention. In the first phase, I introduced the empirical study to participants and obtained their consent to participate in this study by signing the consent form (see Appendix 4). In addition, there were orientations to familiarize experimental groups with treatment procedures and the use of research apparatus. Baseline data and background information were obtained by having three groups of students finish the pre-test and the demographic questionnaire.

One week prior to the intervention, I conducted a pilot study to test the reliability of data collection instruments. Five students from each group were invited to participate in the pilot study and contributed valuable suggestions. Based on their feedback and suggestions, I made three major alterations to the original plan. Firstly, three texts were selected for the intervention, while pilot participants found one of them were too difficult to understand even with the help of multimedia. Therefore, this text was deleted and three sessions were modified into two sessions. Secondly, half of the pilot participants found some new words without annotations that may affect their understanding. Thus, after checking with the teacher, new words in the text were annotated in Chinese and a word list was given in the intervention. Thirdly, participants in the experimental group A complained about the text on the screen was too small to read and focus on. Since it was a technical problem and there was no way to expand the text box, this problem was solved by giving paper text in all groups.

In the second phase, the reading intervention began in the target school. The treatment was offered in two reading sessions within two weeks. As for the treatment procedure (see Figure 6), it was mainly divided into pre-reading, reading and post-reading. Prior to reading, the teacher introduced the reading task and topic for 5 minutes, and students were given 25 minutes to read text and finish a collaborative reading task based on the given materials. During reading sessions, the teacher took the facilitative role to observe learners'

reading and address their doubts in need. Afterwards, the teacher gave corrective feedback of the collaborative task, which helped strengthen students' memorization of the text. The reading session utilized collaborative reading task because of the limited number of research apparatus in the lab, and three students had to share one zSpace computer as a group. The same instructional design was applied to group B and group C to ensure the consistency of treatment.

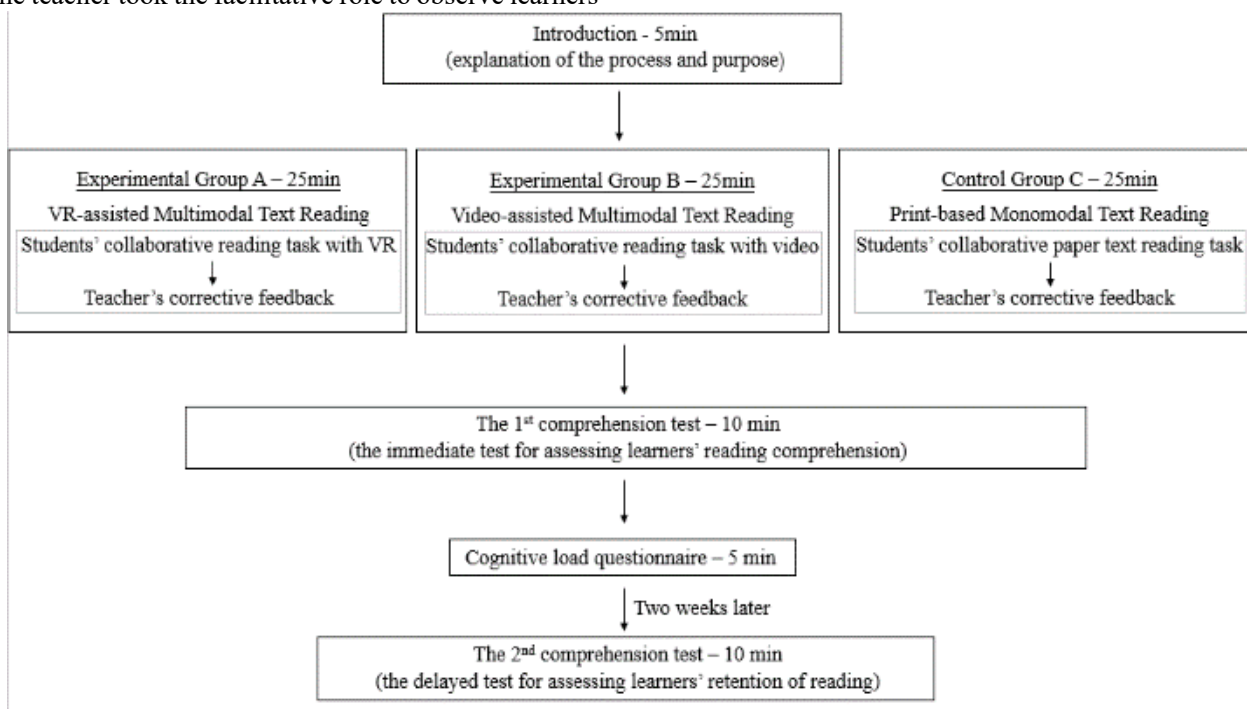


Figure 6. Procedure of the reading treatment

In the third stage, all materials were collected back and the immediate post-test was administered. After finishing the test, students had to complete the cognitive load questionnaire. After each session, students were asked to write a learning journal based on the reading experience. Six students in the two experimental groups were invited on a voluntary basis to participate in a semi-structured interview on the same day, during which they were encouraged to describe their

multimodal reading experience, reflect on the usefulness of multimodal text in comparison to their usual reading practice, and how they applied received multimodal input to answer test questions. Two weeks later, a delayed post-test was administered again among three groups to evaluate the retention effect of multimodal input. Table 3 presented a summary of the data collection procedure.

Table 3. Timeframe of data collection procedure

Time	Stage of plan	Activities	Data set
Week 1	Pre-intervention	Introduction of the study and research apparatus; informed consent form; demographic questionnaire; pilot study	Questionnaires; fieldnotes
		Pre-test to assess prior knowledge	Pre-test scores
Week 2	Intervention	Session 1; learning journal; cognitive load questionnaire;	Post-test scores; questionnaires; journal entries
Week 3	Intervention	Session 2; learning journal; cognitive load questionnaire;	Post-test scores; questionnaires; journal entries
	Post-intervention	Semi-structured interviews	Interview recordings
Week 5	Post-intervention	Delayed post-test	Delayed post-test scores

4.3. Reliability and validity

Since reliability and validity are of paramount importance to research findings, the current research

invested efforts to address the potential threats of both quantitative and qualitative approaches to reliability and validity respectively.

Reliability refers to consistency and replicability

of research findings over time (Nunan, 1992). For the quantitative strand of data, the reliability is concerned with the instruments to measure the effects of multimodal input, such as reading tests and cognitive load questionnaire. As for reading tests, the reading passages and questions utilized in the treatment have been selected from authoritative test papers and checked by two English teachers to ensure that questions in each reading test could identify macrostructural and microstructural reading comprehension. Moreover, the test-retest reliability of the cognitive load scale has been ensured by the Cronbach's alpha value ($\alpha = .87$), indicating satisfactory reliability of items. The implementation of pilot study also reinforced the reliability of reading tests and cognitive load scale by making modifications in tandem with learners' English proficiency.

For the qualitative strand, several approaches have been applied to rule out potential threats of reliability. Firstly, as for the interview instrument, I prepared open-ended questions to elicit learners' recall of the multimodal reading experience and avoided giving personal opinions in case that participants would change opinions due to others' responses in a group interview (Creswell, 2006). In addition, participants were given the right to choose spoken language freely, and all of them chose to use L1 Chinese so that they could share opinions at ease without worrying making grammatical mistakes. In this regard, in-depth and faithful information can be obtained (Bauer, 2000). In terms of journal entries, three leading questions were provided to help learners reflect the multimodal experience and clarify individual cognition (Moon, 1999). The journal entries were not assessed or rated based on a writing rubric but regarded as an approach to understand all participants' perceptions towards the efficacy of multimodal input, which were quantified to generate a coding pattern at the interpretation stage. Peer examination of the categorical matrix was adopted to enhance its reliability. Thus, the reliability of both quantitative and qualitative data collection and analysis was promised.

Validity means "how appropriately and precisely an operationalization matches a construct's theoretical definition" (Mackey & Gass, 2011, p. 204). This study has invested efforts to constitute its internal and external validity. To establish internal validity, the soundness of research design and measuring instruments holds great importance. The potential threat generated by the lack of random sampling in the quasi-experimental research design was addressed by careful selection of three classes from similar average academic background. It is noted that the test-based assessment in multiple-choice format might be criticized because it stands at the behaviorist side to use relatively simple approach to measure learning outcome and cannot capture learners' high-order thinking (Blikstein and Worsley, 2016). However, it is extensively used in educational research due to the high level of objectivity, and the validity of test-based assessment can be strengthened by triangulating data from questionnaires, learning journals and interviews. As for the interview instrument, interviewees' verification and feedback were

gained to ensure respondent validation.

External validity stresses on the generalizability and applicability of research findings to a wider population and learning contexts (Nunan, 1992). The random sampling is the key to generalizing findings to a wider population. However, it is not practical to select the target sample from different schools all around China. Thus, I selected three classes with the average academic performance and language proficiency in Grade 8 of the target school, because they shared similar characteristics with a wider population of Chinese EFL beginners. Although this study focused on the efficacy of multimodal input on the reading aspect of SLA, the research findings and the expanded conceptual framework can shed light on more multimodal learning scenarios so that the generalizability of findings could be achieved.

4.4. Ethical considerations

The empirical study strictly followed the Ethical Guidelines for Educational Research (BERA, 2018) throughout the entire research process, from designing research to conducting fieldwork to reporting findings.

5. Data Analysis

5.1. Quantitative data analysis and statistical results

The quantitative datasets, including demographic questionnaires, three test scores, cognitive load ratings were entered in Statistical Package for Social Sciences (SPSS) version 24.0 to derive descriptive and inferential statistics. The numerical data was analyzed on group basis to capture the overall pattern rather than individual performance. This study set the level of significance at .05 as the criterion for statistical significance since an alpha level of less than .05 is regarded as statistically significant in most educational research.

5.1.1. Effects of different input modalities on learners' reading performance

A 3×3 repeated measures MANOVA was conducted to determine the effect of three input modalities (VR-assisted multimodal text, video-assisted multimodal text, print-based monomodal text) on learners' reading comprehension performance that has been divided into overall comprehension, macrostructural comprehension and microstructural comprehension at three times of testing. There were two independent variables, one is time of assessment as within-subject variable and the other is input modality as between-subjects variable. The pre-test score was utilized as the covariate for excluding any interference from learners' prior knowledge. Before performing statistical tests, assumptions including homogeneity of variance, sphericity and normality have been validated. The justification for using MANOVA was twofold. Firstly, the different levels of reading comprehension act as multiple continuous dependent variables, and using MANOVA instead

of a series of one-at-a-time ANOVAs can reduce the experiment-wise level of Type I error without rejecting true but weak null hypothesis. Additionally, MANOVA can also test if the relationship among the independent variables change over the intervention, and reveal differences not discovered by ANOVA tests.

Table 4 shows the relevant descriptive statistics, including number of participants per group (N), reading test mean scores (M) and standard deviation (SD). There were few differences in learners' pre-test scores, showing that students were at the same baseline level of expository reading ability. In the immediate post-test, participants in the VR group scored the highest on the overall ($M = 3.1190$, $SD = .99271$), macrostructural ($M = 1.3810$, $SD = .66083$) and microstructural reading comprehension ($M = 1.7381$, $SD = .58683$). In the delayed post-test, participants who read VR-assisted multimodal text also performed best on the overall ($M = 2.7857$, $SD = .81258$), macrostructural ($M = 1.1429$, $SD = .41739$) and microstructural reading comprehension ($M = 1.6429$, $SD = .72655$). Participants who received the monomodal print-based text got the lowest scores in both immediate post-test and delayed post-test.

Table 4. Descriptive statistics for reading test score means and standard deviations

	Pre-test Total		Post-test 1 Total		Post-test 2 Total	
	M	SD	M	SD	M	SD
VR ($n = 42$)	1.9286	.99738	3.1190	.99271	2.7857	.81258
Video ($n = 46$)	1.9130	1.13188	2.8043	.98024	2.5870	.95629
Paper ($n = 43$)	1.9302	1.00937	2.4186	1.17984	2.3488	1.34313
Total ($N = 131$)	1.9237	1.04232	2.7786	1.08337	2.5725	1.06721
	Pre-test Macro		Post-test 1 Macro		Post-test 2 Macro	
	M	SD	M	SD	M	SD
VR ($n = 42$)	.7619	.61721	1.3810	.66083	1.1429	.41739
Video ($n = 46$)	.8696	.68666	1.3043	.46522	1.0652	.87945
Paper ($n = 43$)	.7442	.65803	1.0000	.61721	.9070	.78115
Total ($N = 131$)	.7939	.65301	1.2290	.60211	1.0382	.72753
	Pre-test Micro		Post-test 1 Micro		Post-test 2 Micro	
	M	SD	M	SD	M	SD
VR ($n = 42$)	1.1667	.72974	1.7381	.58683	1.6429	.72655
Video ($n = 46$)	1.0435	.66522	1.5000	.80966	1.5217	.78143
Paper ($n = 43$)	1.1860	.58781	1.4186	.90587	1.4419	.79589
Total ($N = 131$)	1.1298	.66097	1.5496	.78656	1.5344	.76759

Figure 7, 8 and 9 display visual representations of how three levels of reading mean scores changed over time for each group. The comparison showed that all groups scored higher in the immediate post-test than pre-test, suggesting that the treatment improved students' reading performance in all conditions. It is noted that experimental group A with assistance of VR achieved higher scores at three levels of reading comprehension in both post-tests than other two groups. Despite a slight decrease, the effect was maintained by three groups at the time of the delayed post-test two weeks later.

Profile Plots

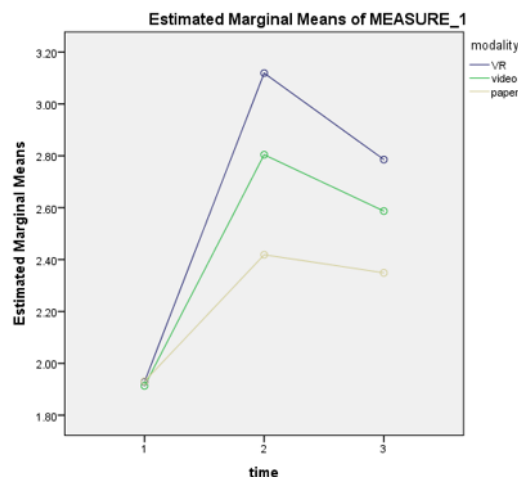


Figure 7. Changes in total reading test scores over time

Profile Plots

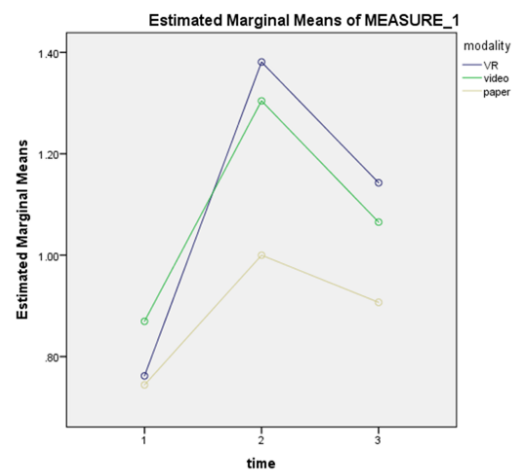


Figure 8. Changes in macrostructural reading test scores over time

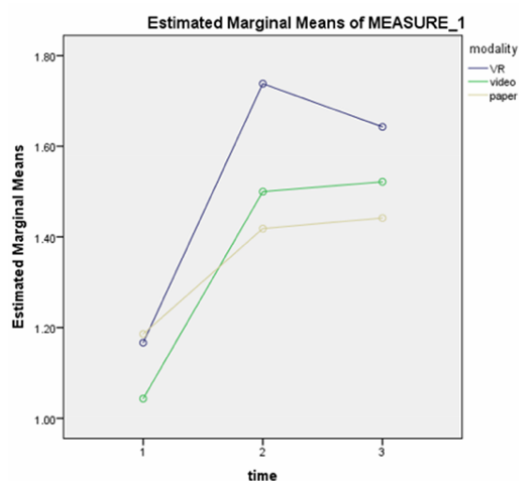


Figure 9. Changes in microstructural reading test scores over time

The results of the MANOVA revealed that the main effect of time within subjects was significant, F

(4, 125) = 15.407, $p < .01$, Wilk's $\Lambda = .670$, partial $\eta^2 = .330$. The main effect of input modality between subjects was significant, $F(4, 254) = 2.738$, $p = .029 < .05$, Wilk's $\Lambda = .919$, partial $\eta^2 = .041$. However, there was no statistically significant interaction effect of time $\times$ modality, $F(8, 250) = .833$, $p = .575 > .05$, Wilk's $\Lambda = .949$, partial $\eta^2 = .026$. One-way ANOVAs were computed to further examine the main effect of modality on short-term and long-term reading performance at overall, macrostructural and microstructural levels between groups.

Tukey's *post hoc* pairwise comparisons were used to identify the significant differences between groups. The results showed that at the time of immediate test, the difference existed between experimental groups and control group, while there was no significant delayed effect of input modality on learners' reading comprehension. In terms of overall reading comprehension, there was significant difference of total reading test score between VR group and paper group, $p = .024 < .05$. As for macrostructural reading comprehension, there were significant differences between VR group and paper group, $p = .033 < .05$, and between video group and paper group, $p = .047 < .05$. In terms of microstructural reading comprehension, no statistically significant difference was found.

To answer the first research question, the results indicated that different input modalities had immediate effect on overall reading comprehension between group A (VR-assisted multimodal text) and group C (print-based monomodal text). Moreover, different input modalities had immediate effect on macrostructural reading comprehension between multimodal text group and monomodal text group. Different input modalities didn't have differential immediate effects on learners' microstructural reading comprehension. In addition, there was no significant delayed effect of input modality on any aspect of reading comprehension.

5.1.2. Effects of different input modalities on learners' cognitive load

To answer the second research question, cognitive load scale was used to assess learners' mental load and mental effort after each session. This study employed Cronbach α to test internal consistency of the cognitive load scale, and the value ($\alpha = .87$) exceeded .80, demonstrating satisfactory reliability of the items. One-way ANOVA was performed to compare learners' cognitive load ratings under three input conditions and examine the effects of multimodal input on learners' mental load and mental effort.

As shown in Table 5, the means and standard deviations of the cognitive load ratings were 2.5274 and .57995 for experimental group A with VR-assisted multimodal text, 2.5671 and .62256 for experimental group B with video-assisted multimodal text, 2.5843 and .72154 for control group C with paper text. There were slight differences of students' cognitive ratings between groups, among which the control group using paper text had the highest mean of cognitive load ratings. The study further compared two components of

cognitive load: mental load and mental effort. For the mental load dimension, the means and standard deviations were 2.1280 and .78287 for the experimental group A, 2.2195 and .69865 for the experimental group B, 2.1570 and .87970 for the control group C. This indicated that video-assisted multimodal text imposed highest amount of mental load on learners than other groups. As for the mental effort dimension, the means and standard deviations were 2.9268 and .68293 for the experimental group A, 2.9146 and .64374 for the experimental group B, and 3.0116 and .72159 for the control group C. Hence, the control group allocated most cognitive capacities in reading the paper text compared with experimental groups.

Table 5. Descriptive statistics for mean scores and standard deviations of cognitive load ratings

	Cognitive load		Mental load		Mental effort	
	M	SD	M	SD	M	SD
VR ($n = 42$)	2.5274	.57995	2.1280	.78287	2.9268	.68293
Video ($n = 46$)	2.5671	.62256	2.2195	.69865	2.9146	.64374
Paper ($n = 43$)	2.5843	.72154	2.1570	.87970	3.0116	.72159
Total ($N = 131$)	2.5600	.64065	2.1680	.78659	2.9520	.68000

The result of one-way ANOVA shown in Table 6 indicated that there were no statistically significant effects of different input modalities on overall cognitive load, mental load and mental effort ($p = .918 > .05$; $p = .867 > .05$; $p = .777 > .05$). It means that participants in three groups had similar levels of cognitive load after the treatment and multimodal text didn't increase germane load or decrease extraneous load compared with monomodal text input.

Table 6. One-way ANOVA on cognitive load ratings

		Sum of Squares	df	Mean Square	F	Sig.
Cognitive load	Between Groups	0.071	2	0.035	0.085	0.918
	Within Groups	50.823	122	0.417		
	Total	50.894	124			
Mental load	Between Groups	0.179	2	0.09	0.143	0.867
	Within Groups	76.543	122	0.627		
	Total	76.722	124			
Mental effort	Between Groups	0.236	2	0.118	0.252	0.777
	Within Groups	57.101	122	0.468		
	Total	57.337	124			

To sum up, the quantitative finding indicated that VR-assisted multimodal text group achieved better overall reading performance than the other two groups, and multimodal text group attained better macrostructural reading comprehension than monomodal text group in the short term, though no statistically significant relationship between input modality and cognitive load was found. This suggested that VR-assisted multimodal text played an important role in fostering L2 learners' overall and macrostructural comprehension ability in the short term without incurring extraneous cognitive load.

5.2. Qualitative data analysis and interpretations

As for qualitative data, content analysis (Garrison, 2006) was performed to probe into participants' reading experience and perceptions by drawing on the

conceptual framework. This current research utilized content analysis because it combines both quantitative (Krippendorff, 2004) and qualitative approaches (Berg, 2001) in alignment with the pragmatist paradigm, and it can be used in an inductive or a deductive way. Another reason for performing content analysis is that it is a particularly useful approach when classifying, summarizing, quantifying and tabulating qualitative data prior to detail explanations. This study used a hybrid process of inductive and deductive approaches to analyze qualitative data, which incorporated both the data-driven inductive approach (Boyatzis, 1998) and the framework-informed deductive approach (Crabtree and Miller, 1999). In this study, the qualitative data analysis was twofold. I used the inductive approach firstly to generate data-driven codes and then I applied deductive approach to generate theory-driven codes, and the two strands of codes were aligned in a systematic way to illustrate learners and teacher's perceptions towards multimodal input. The further analysis began with quantification of qualitative data by using frequency and percentage to show the magnitude of the individual phenomena (Berg, 2001; Morgan, 1993) and reflect the overall tendency, then each coding category would be enriched by in-depth narration.

Following this methodology, I initially familiarized myself with the qualitative data that was extracted from the written feedback and narrative accounts through reflective journals and interviews. The transcription of all interview data was done by a voice-recognition software first and then I checked it to see whether all the information was transcribed accurately by listening to the recording again. Afterwards, I translated the transcripts of interview from Chinese to English. I read through the translated data and obtained a general understanding of the whole pattern. The text segments that answered leading questions clearly were highlighted and coded into positive and negative

comments, and similar comments were further categorized into certain aspect of multimodal input that assisted learners' reading or impeded their reading. Then, in the categorization process, I drew on the conceptual framework to aggregate codes with similar meanings to develop a categorical matrix. The categorical matrix was built on the multimodal input assisted by VR technology in three sensory modalities: graphic and animation (visual), narration and sound effects (auditory), interactivity and manipulative (tactile). After the units of analysis have been identified, I re-read the original text especially the unmarked text to make sure text segments related to the categorized matrix has been covered (Burnard, 1991). Finally, I summarized the frequency and percentage of participants' comments and presented them in tables for comparison. Participants' perceptions of VR-assisted multimodal text and video-assisted multimodal text were discussed separately to provide a holistic understanding of multimodal text utilized in this study.

5.2.1. Learners and teacher's perceptions of VR-assisted multimodal text

The analysis of reflective journals and semi-structured interviews showed that the participants' perceptions towards VR-assisted multimodal text were mostly positive, particularly in terms of visual input and tactile input. Nevertheless, some participants also stated some negative comments towards the multimodal reading experience regarding the time management, complexity of information and the lack of equipment for reading the multimodal text effectively. Table 7 summarizes the categorical matrix of learners' perceptions about VR-assisted multimodal text. Affordances mean what are made possible by the used multimedia while constraints refer to the negative aspects of the multimedia tool that may affect learners' reading comprehension.

Table 7. Categorical matrix of perceptions about VR-assisted multimodal text

	Affordances	%	Example	Constraints	%	Example
Auditory input	Narration	21	I enjoyed listening while reading, because it sounded like a teacher talking to me.	Time cost	12	I think it was a little bit waste of time to listen to the text while reading the text at the same time.
	Sound effect	17	I like the sound of raining and water flowing, because it gave me a real feeling.	Speed	5	The audio recording talked too fast, so I could not follow.
Visual input	Animation	50	VR displayed the animated content vividly, and gives us a sense of immersion.	Complexity	26	The visual content was complex, and it was difficult to find all the details.
	Graphic	43	Some pictures presented the cycle clearly and we can observe things intuitively without imagining it in our mind.	Health concern	14	It hurt my eyes and made me feel dizzy when reading for long time.
Tactile input	Interactivity	45	I can interact with objects in three dimensions to learn more, such as how the chrysalis looks like.	Distraction	17	We may focus on playing with 3D models rather than reading the text.
	Manipulative	31	It gave me a feeling of control, so I can learn at my own pace.	Limited operators	7	I was not the operator, so I didn't feel the interaction with butterfly.

In terms of auditory input, more than one fifth of students found that narration and sound effects aided

their reading comprehension. In the VR-assisted reading treatment, students could hear the recording of the

text to learn pronunciation of target words and had the opportunity of reading with the recording. This made it easier for students to connect the sound with the word and remember it when completing the immediate post-test. The sound effects such as the sound of water flowing to the river and caterpillar eating leaves gave learners a vivid feeling of presence. However, some students commented negatively towards the time cost and fast speed of audio recording, which represented areas for improvement of VR-assisted multimodal reading scenario. One student shared her opinion regarding the auditory input in the focus group interview:

Sara: The sound effects of the reading materials are vivid and attractive, but I don't think it's necessary for our reading because it takes long time to listen to the recording. Also, my classmate wants to hear some paragraphs twice but I think once is enough since it's a reading task not a listening task. It will be more efficient and effective if every student is given a pair of ear-phones and they can listen to whatever they want many times.

When asked what they liked the most about the VR-assisted multimodal text, half of the participants responded that they found the animation of the text content most interesting, such as the growth of butterfly from caterpillar to chrysalis. This sensory stimulation engaged learners in watching the animation and reading the text. In addition, graphics provided static diagrams to facilitate students' overall understanding of the expository text. In other words, visual cues provided participants quick information that can be directly perceived through watching the screen, unlike many non-visual cues such as sound effects of target objects which needs to be learned from students' prior knowledge and other information sources. One participant expressed her affection towards visual input as follows:

Anne: My favorite thing about the VR reading task is the animation that displays the growth of butterfly vividly, and I don't have to imagine it in my mind because watching the animation is sufficient for me to identify different stages of the butterfly's lifecycle. It is said that one image is worth more than a thousand words. However, after the VR-assisted reading session, I think one animation is worth more than a thousand images, because animation is like a thousand pictures displayed at high speed in a series.

Nevertheless, some participants complained about its complexity, which may be explained by the richness of the visual input imposed relatively high level of mental load on learners. In addition, around a quarter of students mentioned the motion sickness and eyes sore they had experienced in the reading process:

Eric: I feel uncomfortable when wearing the glasses and watching 3D objects for more than fifteen minutes, because it makes me feel dizzy and hurts my eyes. I have to take off the glasses and stop reading the text from time to time. I think it has negative influence on my reading test since I could not concentrate on reading the text.

In a similar vein, Ms Li addressed the health concern from the perspective of a recent published educational policy:

Ms Li: Recently a new policy has been

introduced in school to limit students' exposure to electronic products, such as mobile phones, computers and tablets. The VR apparatus, though not mentioned in the policy, is still a kind of computer that may pose detrimental effects on learners' health both physically and psychologically. Students may be addicted to it and become short-sighted easily. Therefore, it is not frequently used in daily courses and we must be very careful when using it in class.

Tactile input is a unique aspect of VR-assisted multimodal text by giving learners' a sense of touch that can be operated by the stylus pen in the air. More than one third of participants stated that they found the VR-assisted multimodal input helpful because they could interact with 3D graphics and control their pace of learning. It shows that VR-assisted multimodal input can be tailored to individual needs and interests. One student shared his related experience in the interview:

Charles: It's amazing! I can drag the butterfly out of the screen and observe it closely by turning it around 360 degrees. You know, in real life, when you get close to a butterfly, it will fly immediately and you can't observe it closely. However, I can catch a butterfly from the screen by using the stylus pen and it won't fly away. I just feel that I can control everything in the virtual world.

Though many students commented it was a positive sensory experience, less than one fifth of students were not fully satisfied with it due to distraction and shortage of equipment. Some students admitted that they spent most time playing with the apparatus rather than reading the text. Ms Li, in the individual interview, also mentioned this constraint according to her observation in class:

Ms Li: Due to the high price of the research apparatus, we can only afford limited numbers in the VR lab. It is not possible for each student to use one VR apparatus, so group work is necessary in the class. I noticed that some group members, if not sit in the middle to operate the apparatus, sometimes engaged in other irrelevant activities. It is difficult for a me to supervise 9 groups of students simultaneously, and the effective implementation of VR-assisted lessons is largely depended on their self-discipline. I think tactile input is a key feature of VR, but it needs to be utilized more effectively by students in class. It may take some time because students currently are more interested in the instrument itself than the knowledge presented.

In summary, participants found reading VR-assisted multimodal text interesting and helpful because narration and sound effects from auditory input, animation and graphics from visual input and interaction and manipulation from tactile input facilitated their understanding of expository texts. Despite the general positive attitude of participants, several problems such as time cost, fast speed, distraction concerning three modes of input were pointed out by some students and the teacher, which may help to explain why there was no significant effect of VR-assisted multimodal text on learners' retention after two weeks.

5.2.2. Learners and teacher's perceptions of video-assisted multimodal text

Compared with VR-assisted multimodal text,

video-assisted multimodal text relies on auditory and visual stimulus. The same categorical matrix for auditory and visual input has been applied to analyze

participants' perceptions about video-assisted multimodal text (see Table 8).

Table 8. Categorical matrix of perceptions about video-assisted multimodal text

	Affordances	%	Example	Constraints	%	Example
Auditory input	Narration	28	I like watching the video because the character explained the water cycle clearly.	Time cost	6	Sometimes I want to switch it back and listen to one part but it took me a lot of time to start it again.
	Sound effect	11	The sound of butterfly's growth was very vivid, and it helped me remember the process.	Speed	48	The character in the video talked too fast, and I could not follow and take notes.
Visual input	Animation	30	The video showed the water cycle in four animated steps, and I can remember how water changes.	Complexity	15	The video content was a little complicated, and I found it hard to understand in English.
	Graphic	4	I like the last scene of the video, because it summarized all the information in the text and helped me answer questions.	Health concern	0	No comment.

As for auditory input, a number of participants found watching the video interesting because of vivid narration and sound effects. There was a cartoon character who explained each stage of the expository text clearly, and students regarded the character as a peer to learn from. Background music and sound effects also engaged students in reading the multimodal text. However, nearly half of the students claimed that it talked too fast and they could not control the speed of the video to slow it down. In addition, it was difficult to switch the video back to a certain part and it took long time to watch it from the beginning again. Ms Li shared her similar opinion towards the video-assisted multimodal text as follows:

Ms Li: I often use video in class as an introduction, aiming to stimulate students' interests rather than give them a task. Therefore, when students need to answer questions based on the video, they may pay more attention to what it talks about and find that they can get a general idea but it is too fast for them to write down notes in detail. Students like the narrator probably because it is a popular cartoon character and they would be more focused when listening to it than listening to me.

Regarding visual input, participants found the colorful and animated display shown in the video could be a great aid for reading comprehension. The majority of information was presented in animation while there was a summarized figure at the end of the video. 30% of students found the animation helpful because it illustrated the whole process of butterfly's growth and water's journey vividly and coherently. However, 15% of students felt overwhelmed because the video contained too much information and the subtitle was in English rather than Chinese. One student described her confusion in detail:

Lara: I think the video is interesting and visually attractive, but I still find it hard to understand because the cartoon character talks in English and the subtitle is in English, and it takes me a while to translate it in my mind, but when I finally understand one sentence, the video has already progressed to next stage.

Especially I have no prior knowledge of water cycle, so I think it is too difficult for me to understand the video.

In contrast to VR-assisted multimodal text, there was no comment of health concern, indicating that video is a widely accepted multimedia tool in the classroom setting and students feel comfortable with it. To sum up, the majority of participants found video-assisted multimodal text aided their comprehension because of cartoon character's narration, vivid sound effects, comprehensive animations and graphics, while some students reported problems such as time cost, fast speed and complex content that need to be tackled through careful selection of videos in accordance to students' level of language proficiency.

To answer the third research question, learners mainly held positive attitude towards multimodal input in the reading session, and they found the multimodal text assisted by VR and video interesting and effective in helping them understand the expository text because of multimedia aids in different modalities. It is also noted that some technical problems constrained students from reading effectively and need to be addressed in the future implementation of multimodal text reading.

6. Discussion

Based on the research findings, this study argued that VR-assisted multimodal input facilitated Chinese 8th grade EFL learners' overall and macrostructural reading comprehension in the short term without incurring extraneous cognitive load.

6.1. The effects of input modalities on learners' reading comprehension

The experimental results showed that VR-assisted multimodal input significantly improved Chinese 8th grade EFL learners' overall and macrostructural reading comprehension in the short term. This supported Jones and Plass' (2002) assumption that "pictorial

information provided in addition to text may help support macro-level processing in L2 computer-based reading activities” (p. 548). The findings of the present study corroborated Al-Seghayer’s (2007) research that the use of structural devices improved learners’ ability to identify main ideas and construct appropriate mental representations of an electronic text. The positive effect of multimodal input on learners’ macrostructural processing was also reported by Abdi (2012), which demonstrated that the exposure of electronic texts supported readers’ macrostructural construction and organization. Macrostructural comprehension entails two levels of processing, one is selecting important textual information from individual units (construction), and the other is connecting selected information into a coherent mental representation (organization). The positive effects of VR-assisted multimodal input on learners’ macrostructural comprehension can be explained by effective monitoring on the two levels of processing. Firstly, the visual support especially the animated feature of VR-assisted multimodal input was effective in introducing thematic units, clarifying complex concepts into simple visual display and providing a holistic understanding of discourse organization. Secondly, the tactile input of VR-assisted multimodal text made it possible for learners to see objects in three dimensions and construct coherent representation of each stage in the butterfly growth or water cycle in a unified form. Thus, participants in the VR-assisted multimodal text group were able to identify individual units, recognize the interrelations and integrate them into coherent mental model, thereby achieving high level of macrostructural comprehension.

As for microstructural comprehension, there were no significant differences between multimodal input and monomodal input, indicating that participants who read the multimodal text presented in audio, visual and tactile modes did not remember more words and textual details than participants who read paper text. Similar findings were also found in Ariew and Ercetin’s (2004) research, stating that there was no causal relationship between multimedia-assisted annotations and learners’ microstructural reading comprehension. This study went beyond multimedia-assisted annotations and included multimodal presentation of the whole text that has not been examined in pervious literature, and the result can be explained by the constraints of multimedia technology and specific reading context in this study. Based on participants’ narrative accounts, complex content shown in limited time diverted their attention from certain details and affected their memorization of textual information, although the detail information in the text has been reinforced in different modes of input. In addition, the initial ‘wow’ effect brought by VR technology could be translated into more attention on the multimedia itself rather than the reading content and language, thereby distracting learners from concentrating on the details illustrated in the text. It is also noted that the study only focused on expository reading in a CLIL context.

It is also worth noticing that there was no

significant effect of input modality on learners’ retention of text, which corroborated previous research findings (Brett, 2001; deHaan, 2010; Moreno, 2002) that certain foreign language multimedia learning environment may not affect learners’ language retention in the long term. One possible explanation is that students were engaged in the multimodal reading session during the initial learning phase, while with the diminishing ‘wow’ effect they are less likely to retrieve newly acquired information to foster long-term learning (Roediger & Karpicke, 2006). Another possible explanation is the lack of incentive for learners to complete the delayed post-test given that they already had finished two similar print-based tests two weeks ago, and the negative testing effect may influence their reading performance.

However, the results should be interpreted with caution and situated in the specific research context. The simple test-based accountability may not be able to generate accurate estimates of gains in student performance. Although these scores provide useful information that contributes to students’ reading growth, looking at reading test scores only would silence other valuable indicators and bias the evaluation of the multimodal reading intervention. In this study, the reading test was formatted in multiple-choice and blank-filling questions for the objective grading, while students could answer the questions with lucky guess or draw on problem-solving strategies to complete the reading test without retrieving newly acquired knowledge, and this might be partially responsible for the absence of significant treatment effects on learners’ retention. Moreover, standardized exams using limited number of closed questions leave little space for learners to display their high-order thinking, such as analysis, evaluation and creativity. Given the limited scope of expository texts, such influence is disproportionate to any intrinsic value they may have on educational outcomes. Furthermore, the conventional paper-based test used in this study was not aligned with the different modalities of text input in the treatment. Students who read multimodal text in the treatment did not finish the post-test in the same format due to technological limitation and complexity of collecting answers digitally. The mismatch between intervention and assessment is likely to affect the research findings. In addition, students may have negative feelings towards test-based assessment due to the stress, previous failed experience and frequent testing, which could lead to demotivation in completing post-tests. In other words, the objective measure of learners’ reading performance could not fully capture the effectiveness of multimodal input. Therefore, it is necessary to combine them with other subjective measures, such as qualitative data collection methods used in this study to fully capture students’ reading development in the multimodal learning environment.

6.2. The effects of multimodal input on learners’ cognitive load

An unexpected finding of the study was that the learners’ overall, extraneous and germane cognitive

load did not show significant differences when reading, watching and interacting with multimodal text, compared with traditional print-based text reading. This means that neither the modality effect nor the redundancy effect was observed based on learners' self-reported cognitive load ratings in the research. The result corroborated few studies since the major component of literature situated findings in either modality or redundancy effect without the third possible result such as no effect. One explanation of this surprising 'no-effect' finding is the pervasiveness of multimodal literacy in the digital era, since students live in a highly visual world and they are exposed to a multimodal environment both in print and on screen. As the multimodal text becomes the new norm, students may find that reading screen-based text does not require more attention and processing compared with print-based text. It is also noted that the complex content and difficult words in the expository reading reported by some participants did not cause learners' cognitive system to be overloaded. The overwhelming demands of cognitive processing in reading relatively complex expository texts were offset by segmentation, which means dividing the passage into learner-paced segments and allowing learners to fully understand each part of the presentation before moving to the next part by clicking the 'continue' button on the screen. The self-controlled input presentation also reduced students' representational holding at one time and they can process information at their own pace. This segmentation principle and pacing principle underlying the cognitive theory of multimedia learning have been validated by multiple studies (Aly, Elen, & Willems, 2005; Lusk, 2008; Mayer, 1999; Mayer & Chandler, 2001). Thus, the positive effects of text segmentation counteracted the increased mental demand of processing complex content, leading to no statistically significant effect of multimodal input on learners' cognitive load.

Although the present study showed that multimodal input did not increase or decrease learner's cognitive load to a large extent, it is premature to conclude that multimodal text assisted by VR and video had no effect on learners' cognitive load. The rationale is two-fold. To begin with, this study approached cognitive load by adapting a self-reported scale (Paas, 1992), the validity of which has been confirmed in multiple studies (Szulewski et al. 2018; van Gog & Paas, 2008). The clarity and sensitivity of this subjective technique lend itself to be an extensively used method while at the same time the self-reported nature is regarded as its major weakness. Antonenko & Niederhauser (2010) suggested that cognitive load should be regarded as a dynamic process and the EEG-based physiological measure should be used to measure it to provide a more comprehensive picture than the self-reported scale. Although the present research specified the subjective rating scale into mental effort and mental load and examined the effects of multimodal input on the two constructs respectively, more objective measures can be applied to strengthen the reliability of research findings. Furthermore, the EFL learners in the current research

were only engaged in two reading sessions, and the exposure to multimodal text was far from enough to draw a definite conclusion regarding the effects of multimodal input on learners' cognitive load. Thus, this study inquired into participants' perceptions towards the efficacy of multimodal text to understand learners' cognitive processing from an inner perspective.

6.3. Learners' and teacher's perceptions about the efficacy of multimodal text

In general, learners held positive attitudes towards the effectiveness of multimodal text and found multimodal input aided their comprehension. Contextualized images and animation were the most appreciated features of multimodal text, found useful by half of learners in both experimental groups. The significant role of visualization in scientific reading has also been addressed by Mason, Tornatora and Pluchino (2013), because they made complex processes visible and helped readers construct mental representations. Based on journal entries and interview data, the majority of learners believed that VR-assisted multimodal input was more effective in assisting their L2 reading than video-assisted multimodal input and traditional print-based monomodal input that they often receive in daily practice. The main reason lies in the unique exposure of tactile input, which gave learners a sense of immersion that cannot be experienced in other text input. Positive perceptions towards the tactile input were also found in Limniou et al.'s (2008) research that showed 3D immersive VR-assisted learning environment elevated learners' interest and motivation compared with learning in a 2D animated environment. In the VR-assisted reading scenario, learners were exposed to a simulated version of the reality which may not available or possible in the brick-and-mortar classroom. According to learners' narration, the interactive affordance of VR technology allowed them to experience, establish kinesthetic relationships with the virtual world and receive feedback contingent on learners' responses (Moreno et al. 2001), and the manipulative affordance allowed them to take up the active role to learn at their own pace. In this light, the tactile input enabled readers to construct haptic model of the text and place themselves in the reality that has been brought from outside to inside the classroom (Evans and Green, 2006).

Nevertheless, some negative aspects of multimodal input have been pointed out in learners' journal entries and the teacher's observation, which constrained readers to construct mental models effectively and prompted practitioners to reflect the potential disadvantages of the multimodal approach in the CLIL reading context. For both experimental groups with multimodal text input, long time of watching electronic screen that was filled with text, sound, graphic and animation was harmful to learners' eyes, especially the immersive nature of VR technology made them feel dizzy after long periods of interacting with the virtual world. Moreover, some students felt lost and overwhelmed in the face of so much information shown in a fast pace. Distraction has also been found to be a

factor in technology-enhanced learning environment when students failed to engage in robust learning (Greene, Moos, & Azevedo, 2007). These negative comments were also reported in previous studies in multimedia-assisted language learning (Lu, 2008; Thornton & Houser, 2005; Wang & Higgins, 2006). These constraints explain some participants' negative comments towards the multimodal input, and these should be taken into consideration along with positive effects of multimodal input on L2 reading in the pedagogical design.

6.4. Theoretical implications for SLA research

According to Mayer (2008), theory and practice are actively engaged in dialog, and this dialog can be built when there is a “two-way street between cognitive science and instruction” (p. 760). This means there is a reciprocal relationship between learning theory and educational practice in which the learning theory lays theoretical understanding for the educational practice whereas the instructional practice is designed on the basis of theoretical framework and further inform the theoretical development. This study is such a dialog that allows for interaction between theory and practice. Theoretically, the present study synthesized two theoretical perspectives and brought forward an integrated framework of cognitive theory of learning with VR. Grounded in the integrated framework, this study tested the efficacy of multimodal input on Chinese EFL learners' reading comprehension by collecting and analyzing quantitative experimental results and qualitative interpretations. The overall research findings supported the tenets of both input hypothesis and the cognitive theory of multimedia learning that the provision of comprehensible input through multiple modalities can facilitate learners' information processing and improve their reading performance in the short term. I want to use this research as a platform to inform the SLA field by problematizing underlying assumptions of the conceptual framework.

Based on the findings of the multimodal text reading practice, the present research can inform the theoretical underpinnings of multimodal learning in cognitive account of SLA. The central issue of this framework is how multimodal input helps people process information and achieve better learning outcome, and this study specifically revealed how VR-assisted multimodal text fostered learners' reading comprehension and validated the framework. From a holistic perspective, the present research focused on multimodal input provision and learners' cognitive processing of input to generating learning output. This study approached the two theoretical perspectives in SLA and multimedia learning in light of multimodality.

Firstly, the current research opened the theoretical lens to encompass multimodal input and examined the way in which VR technology enhances input in a multimedia learning environment. Instead of limiting the scope in providing multimodal input, this study has fully exploited multiple affordances of VR technology in providing auditory, visual and tactile input and

explained how the different modes of input entered in learners' working memory. It is worth noting that the constraints of multimodal input were also revealed in the current research, suggesting that negative aspects of multimodal input should also be a focus of inquiry to obtain a comprehensive understanding towards the efficacy of multimodal input. In this light, this research has taken the field beyond Krashen's theory of comprehensive input to an understanding that how learners process the multimodal input and construct different mental models into coherent understanding. This study also suggested that SLA research with a focus on linguistic input should consider the way how technology changes linguistic input and how learners' access to different affordances of multimodal input might affect acquisition.

Secondly, this study expanded the three assumptions (the dual-channel assumption, the limited capacity assumption, the active processing assumption) underlying the cognitive theory of multimedia learning (Mayer, 2005) for further application of this framework in SLA field. Incorporated VR technology into the multimodal learning environment, this study challenged the dual channels assumption and suggested that a third modality could be added to the dual coding hypothesis (Paivio, 1986) and cognitive theory of multimedia learning (Mayer, 2005) since the tactile input is not provided in purely auditory and visual modalities. Built on the triple-channel hypothesis, this study posited that learners could develop an additional haptic mental model of reading materials and improve learning outcome as a result of more thorough processing. The facilitative effect of VR-assisted multimodal input on learners' objective macrostructural comprehension performance validated the triple-channel hypothesis and emphasized the instrumental role of tactile input in assisting learners identify different units of information and construct coherent mental representation. The research findings also indicated that the additional dimension of input modality did not impose cognitive load or alleviate cognitive load. Thus, this study added another possible result of the limited capacity assumption that do not fall into redundancy effect or modality effect, which requires further evidence to measure effects on working memory load in a theory-related manner. Although there was no statistically significant effect of multimodal input on learners' cognitive load, learners reported high level of interest and motivation and positive attitudes towards the effectiveness of multimodal text. Therefore, affective and motivational aspects of multimedia learning can be added to the active processing assumption, which involves not only the integration of prior knowledge with mental models but also the affective domain that may influence acquisition. Together the three extended assumptions form the cognitive basis for multimodal learning and provide a start point for designing multimodal instruction.

In summary, this empirical study expanded the theoretical landscape by situating multimodality in SLA through a discussion of input hypothesis (Krashen, 1985) as well as cognitive theory of multimedia

learning (Mayer, 2005) in a VR-assisted multimodal learning environment. Guided by the integrated framework, the mixed research findings from objective assessment and subjective narration further extended the scope of comprehensible input and the three assumptions underlying the cognitive theory. The interplay of theory and practice shown in this study provides a solid conceptual ground and empirical evidence for future investigation of multimodality in the field of SLA.

7. Conclusion

7.1. Main findings

This study argued that VR-assisted multimodal input significantly improved Chinese 8th grade EFL learners' overall and macrostructural reading comprehension in the short term without incurring extraneous cognitive load. Situated in the integrated framework of cognitive theory of multimedia learning with VR, this study examined the effects of VR-assisted multimodal input on Chinese 8th grade EFL learners' reading comprehension by triangulating objective reading test performance, cognitive load ratings and participants' narrative accounts to obtain an overall understanding of multimodal input in breadth and depth. Based on the research findings, there were statistically significant differences in the short-term reading performance at the level of macrostructure, with superior performance of the students receiving multimodal text treatment. Besides it, there was no effect of multimodal text on students' short-term microstructural reading comprehension and long-term retention of the text. In addition, there were no differences in terms of learners' cognitive load, indicating that multimodal text did not incur extraneous load or increase germane load. Participants' positive feedback displayed the affordances of VR in assisting their expository reading, while the negative comments and suggestions required researchers, teachers and technological designers to further improve multimodal input in accordance with students' cognitive capacity in the multimodal learning environment.

7.2. Implications

Implications for teaching with advanced technology fall into the interdependent categories of materials design and student training. In terms of materials design, teachers need to select appropriate 3D models from the database and incorporate flexibility in teaching materials and assignments so that students can choose among a variety of tools or strategies in order to customize learning in accordance to their levels of language proficiency. Teachers can develop a repertoire of instructional approaches to encourage learners to construct multimodal representation and process information actively. Moreover, teachers need to provide training to orient students to the multimodal text input, otherwise readers may wander at random through multimodal text and not able to construct coherent mental representation among text units. Therefore, teachers

need to acquire technology-supported skills and pedagogical knowledge to integrate advanced technology with language teaching.

As for technological implications, the constraints mentioned by participants in this study require further improvement of state-of-the-art infrastructure and facilities. Technology providers in the education field need to take pedagogical design into consideration and offer practitioners enough training and technical support to facilitate multimedia assisted instruction and learning. In addition, this study was conducted in a context of governmental support in promoting VR technology in education. The mixed findings remind policy makers to think twice before implement VR technology in the education field to a large extent and treat the immediate 'wow' factor with caution. Despite the benefits of VR, there are still some challenges that should be addressed before implementing at large scale.

7.3. Limitations and future directions

This exploratory research took an important step in integrating VR technology with EFL learners' expository reading, while its generalizability was reduced due to the non-random sampling of participants, short time frame, nature of expository reading and the formal school context. The unanimous background of participants as 8th grade Chinese EFL learners narrowed the generalizability of findings. Moreover, the study was carried out over a short time frame with only two reading sessions. Therefore, participants had limited exposure to multimodal text, and this study only captured learners' initial excitement towards new form of technology but may ignore their decreasing motivation and affection that influence their learning outcome in the long-term. Future studies can take a longer period of time for the treatment and observe learners' reading trajectory in dealing with multimodal text. Moreover, this study focused on expository reading and situated the treatment in a CLIL context. Different types of reading and other facets of language learning can be the focus of inquiry in future studies. Lastly, this study focused on learning in the school context and utilized conventional assessment tools. There has been very little empirical research on the instructional value of multimodal language learning in informal settings, and it is recommended that researchers deploy advanced multimedia technology to assist students' informal language learning.

References

- Altarriba, J., Knickerbocker, H., Trofimovich, P., & McDonough, K. (2011). Acquiring second language vocabulary through the use of images and words. Applying priming methods to L2 learning, teaching and research, 21-48.
<https://doi.org/10.1075/llt.30.06alt>
- Antonenko, P. & Niederhauser, D. (2010). The influence of leads on cognitive load and learning in a hypertext-assisted learning environment. *Computers in Human Behavior*, 26(2), 140-150.
<https://doi.org/10.1016/j.chb.2009.10.014>

- Atif, A., Richards, D., & Bilgin, A. (2013). A student retention model: empirical, theoretical and pragmatic considerations. In 24th Australasian Conference on Information Systems (ACIS) (pp. 1-11). RMIT University.
- Azevedo, R., Greene, J. A., & Moos, D. C. (2007). The effect of a human agent's external regulation up-on college students' hypermedia learning. *Meta-cognition & Learning*, 2, 67- 87.
<https://doi.org/10.1007/s11409-007-9014-9>
- Berg, B. L. (2007). *Qualitative research methods for the social sciences*. London: Pearson.
- Brinton, L. J. (1991, August). The origin and development of quasimodal 'have to' in English. In Workshop on the Origin and Development of Verbal Periphrases, 10th International Conference on Historical Linguistics (ICHL 10), Amsterdam.
- Chow, A., Andrews, S., & Trueman, R. (2007). A 'Second Life': can this online, virtual reality world be used to increase the overall quality of learning and instruction in graduate distance learning programs. In Proceedings of the Association for Educational Communications and Technology International Convention (Vol. 2, pp. 75-83). Association for Educational Communications and Technology, Bloomington, IN.
- Clark, R. E. (2001). *Learning from media*. Greenwich, CT: Information Age Publishing.
- Cohen, L., Manion, L., Morrison, K., & Morrison, R. B. (2007). *Research methods in education*. Routledge.
<https://doi.org/10.4324/9780203029053>
- Crabtree, B. F. & Miller, W. L. (1999). *Doing Qualitative Research*. London: Sage
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, 209, 240.
- DeHaan, J., Reed, W. M., & Kuwanda, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, 14(2), 74-94.
<https://doi.org/10.9762/digraj.4.2> 47
- Dewey, J. (1998). *The essential Dewey: Pragmatism, education, democracy* (Vol. 1). Indiana University Press.
- Duffy, J. (1985). Effects of instruction in noting and using an author's pattern of writing on sixth graders' understanding and memory of expository text (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.
- Ellis, N. (1999). Cognitive approaches to SLA. *Annual Review of Applied Linguistics*, 19, 22-42.
<https://doi.org/10.1017/S0267190599190020>
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, 5(1), 80-92.
<https://doi.org/10.1177/160940690600500107>
- Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education*, 9(1), 1-8.
<https://doi.org/10.1016/j.iheduc.2005.11.001>
- Gass, S. (1997). *Input, Interaction, and the Second Language Learner*. Mahwah, NJ: L. Erlbaum.
- Gottlieb, M. (1999). *The language proficiency handbook: A practitioner's guide to instructional assessment*. Springfield, IL: Illinois State Board of Education.
- Grundhauser, E. (2016). Did a Silent Film About a Train Really Cause Audiences to Stampede? *Atlas Obscura*.
- Hitchcock, G., & Hughes, D. (1995). *Research and the teacher: A qualitative introduction to school-based research*. Psychology Press.
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121-130.
<https://doi.org/10.1016/j.compedu.2013.07.008>
- Jamaludin, A., Ho, C. M. L., & Chee, Y. S. (2007). Argument-Based Negotiation and Conflict Resolution through Enactive Role Play in Second Life. *Frontiers in Artificial Intelligence and Applications*, 162, 561.
- Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, 86(4), 546-561.
<https://doi.org/10.1111/1540-4781.00160>
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Prentice Hall.
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Harlow: Longman.
- Lan, Y. F., & Sie, Y. S. (2010). Using RSS to support mobile learning based on media richness theory. *Computers & Education*, 55(2), 723-732.
<https://doi.org/10.1016/j.compedu.2010.03.005>
- Legros, D., Crinon, J. (2002). *Psychologie des apprentissages et multimedia*. Armand Colin, Paris.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. *Handbook of second language acquisition*, 413-468. New York: Academic Press.
<https://doi.org/10.1016/B978-012589042-7/50015-3>
- Lorenz, B., Green, T., & Brown, A. (2009). Using multimedia graphic organizer software in the pre-writing activities of primary school students: What are the benefits?. *Computers in the Schools*, 26(2), 115-129.
<https://doi.org/10.1080/07380560902906054>
- Mackey, A., & Gass, S. M. (2011). *Research methods in second language acquisition: A practical guide*.

- John Wiley & Sons.
<https://doi.org/10.1002/9781444347340>
- Mackey, A., & Gass, S. M. (2013). Stimulated recall methodology in second language research. Routledge.
<https://doi.org/10.4324/9781410612564>
- Mangen, A., Walgermo, B., & Bronnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61-68. <https://doi.org/10.1016/j.ijer.2012.12.002>
- Mason, L., Tornatora, M. C., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing*, 28, 851-872.
<https://doi.org/10.1007/s11145-015-9552-5>
- Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. 41, pp. 85-139). Academic Press.
[https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1), 43-52.
https://doi.org/10.1207/S15326985EP3801_6
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning*, 43.
<https://doi.org/10.1017/CBO9780511816819.004>
- Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760-9. <https://doi.org/10.1037/0003-066X.63.8.760>
- McCormick, S., & Zutell, J. (2015). *Instructing students who have literacy problems* (7th ed.). Boston, MA: Allyn & Bacon.
- Moon, J. (1999). *A handbook of reflective and experiential learning*. London: Routledge.
- Moreno, R. (2006). Learning in high-tech and multimedia environments. *Current directions in psychological science*, 15(2), 63-67.
<https://doi.org/10.1111/j.0963-7214.2006.00408.x>
- Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry*, 20(8), 1045-1053. <https://doi.org/10.1177/1077800413513733>
- Neo, M., & Neo, T.-K. (2009). Engaging students in multimedia-mediated Constructivist learning – Students' perceptions. *Educational Technology & Society*, 12 (2), 254-266.
- Nunan, D. (1992). *Research methods in language learning*. Cambridge University Press.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGrawHill.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of educational psychology*, 84(4), 429.
<https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63-71.
https://doi.org/10.1207/S15326985EP3801_8
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, 19(2), 221-243.
[https://doi.org/10.1016/S0747-5632\(02\)00015-8](https://doi.org/10.1016/S0747-5632(02)00015-8)
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249-255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Roehling, J. V., Hebert, M., Nelson, J. R., & Bohaty, J. J. (2017). Text structure strategies for improving expository reading comprehension. *The Reading Teacher*, 71(1), 71-82.
<https://doi.org/10.1002/trtr.1590>
- Rost, M. (2002). *Teaching and Researching Listening*. Longman, London.
- Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.
- Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 22(3), 375-387. <https://doi.org/10.1002/acp.1413>
- Son, J. B. (2003). A hypertext approach to foreign language reading. *Australian Review of Applied Linguistics. Supplement Series*, 17(1), 91-110.
<https://doi.org/10.1075/araalss.17.07son>
- Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, 10(3), 251-296.
<https://doi.org/10.1023/A:1022193728205>
- Tomasio, D., Chang, L., Caparellia, E. C., & Ernst T. (2007). Different activation patterns for working memory load and visual attention load. *Brain Research*, 1132, 158-165.
<https://doi.org/10.1016/j.brainres.2006.11.030>
- Walsh, M., Asha, J., & Spranger, N. (2007). Reading digital texts. *The Australian Journal of Language and Literacy*, 30(1), 40.
- Yvonne, M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1), 6-16.
<https://doi.org/10.1177/1558689809349691>

Call for Papers

Submit via <https://jlt.ac/>

Areas of Interest:

Language teaching intervention and experiments; Curriculum development; Language teacher education; Bilingual education; New technologies in language teaching; Testing, assessment, and evaluation; Educational psychology, and more.

We accept the following types of submission:

1. Research article: (6,000 to 8,000 words)
2. Review: (3,000 to 8,000 words)
3. Book review: (up to 3,000 words)
4. Features: (3,000 to 8,000 words)

Scan to submit your articles* &
read more articles for free.

*Article Processing Charges Apply.



Contact: editor@jlt.ac



Pioneer
Publications

ISSN (Online)
2770-4602