International Journal of Learning, Teaching and Educational Research
Vol. 18, No. 1, pp. 76-91, January 2019
https://doi.org/10.26803/ijlter.18.1.6

Multimedia Learning: Contributions of Learners' Verbal Abilities and Presentation Modalities

Sunjung Kim
University of Central Arkansas
Department of Communication Sciences and Disorders

Linda J. Lombardino
Special Education, School Psychology, and Early Childhood Studies
College of Education
University of Florida

Abstract. In the current investigation, two studies were designed to examine the effects of learner's individual cognitive differences and presentation types on students' learning, using an experimental research design. Prior to the participation in one of two studies, college students were tested on measures of verbal ability, visual memory, and background knowledge. In the first study, 114 students were presented with narration-based lessons (narration + image or narration-only) while in the second study, 190 students were presented with text-based lessons (text + image, text + image + narration, or text-only) followed by comprehension questions. In both studies, verbal ability was a strong concurrent predictor of learning outcomes irrespective of the type of instructional media. Behavioral and eye-gaze data indicated that multimedia presentations resulted in better learning outcome than single media presentations both in the narration-based and text-based conditions and that redundant presentations of information did not improve learning. Findings support the use of multimedia instructional platforms in conjunction with strengthening students' verbal skills.

Keywords: multimedia; verbal ability; retention; transfer knowledge; eye gaze.

1. Introduction

Multimedia learning becomes increasingly available and accessible due to advances in educational technology. 97% public school teachers reported to have at least one computer in their classrooms with an average student to computer ratio in the classroom is 5.3 to 1 (Gary, Thomas, & Lewis, 2010). In order to facilitate learning effectiveness, we aimed to examine the impact of individual cognitive differences and research-based instructional methods on the learning outcomes of college students in two studies. The conceptual framework for this study was grounded in Mayer's cognitive theory of multimedia learning (Mayer, 2002, 2014) and Paivio's dual coding theory (Paivio, 1990; Sadoski & Paivio, 2013).

Mayer's (R. C. Clark & Mayer, 2016; Mayer, 2009) cognitive theory of multimedia learning posits that a robust understanding of how people learn is the basic foundation of instructional design and learning technology. The most fundamental tenet of Mayer's work is that instructional methods (e.g., ways of presenting materials) can positively or negatively affect learning. On principle, learners learn better in multimedia than in a single media contexts. This *multimedia principle* has been applied to various studies of academic subjects including language (Takacs, Swart, & Bus, 2015 for review), mathematics (Chiu & Churchill, 2015), and science (Mason, Pluchino, Tornatora, & Ariasi, 2013). Paivio's dual coding theory (J. M. Clark & Paivio, 1991; Paivio, 2014) posits that humans process information visually and verbally through separate but interrelated channels. Dual coding theory was first applied to the cognitive domain of memory (e.g., Jessen et al., 2000) and then expanded to other areas. For example, concurrent activation of verbal and visual information has been shown to enhance reading comprehension (Sadoski, Goetz, & Fritz, 1993), written composition (Goetz, Sadoski, Stricker, White, & Wang, 2007), second language learning (Jared, Poh, & Paivio, 2013), patient education (Goolsby & Sadoski, 2013), and advertising effectiveness (P. C. Lin & Yang, 2010).

Research shows the positive effect of multimedia platform on students' learning but it is not clear in prior literature how students' individual differences affect learning in the multimedia environment. As posited by the dual coding theory, students' high or low ability to process visual and verbal information can affect learning outcomes. For example, students with strong visual abilities understand learning materials better and remember more study content than students with weaker visual abilities (e.g., Brunyé, Taylor, & Rapp, 2008). Students' verbal ability is also likely to play a powerful role in acquiring new knowledge during learning (Pazzaglia, Toso, & Cacciamani, 2008). In a study of college students who were learning a second language through a multimedia (pictorial and written annotations) teaching module, Jones (2009) reported that students with high verbal ability performed better on vocabulary and story recall tests than students with low verbal ability.

With the advances in educational technology, we were interested in how individual learner's cognitive characteristics and research-based learning principles would work together in the context of multimedia instruction. We conducted two investigations to address the effects of these variables in different media contexts. In the first study, we investigated the extent to which individual differences and instructional design affect learning in narration-based lessons (*narration* + *image* and *narration-only*). In the second, we investigated the impact of individual differences and instructional design on learning in text-based lessons (*text* + *image*, *text* + *image* + *narration,* and *text-only*). In both studies, verbal ability, visual memory, and background knowledge served as independent variables representing individual differences among learners. The presentation conditions served as independent variable that represented instructional design features. Based on the previous research (Mayer, 2002, 2014), we hypothesized that multimedia instruction would improve students' learning compared to single media instruction regardless of the presentation modalities (text-based or narration-based) while redundant information would not increase students' learning outcomes. Further, we hypothesized that individual difference would affect learning outcomes even after controlling for instructional design and background knowledge.

2. Methodology

Quantitative methods were used to address the study hypotheses regarding the effect of individual differences and instructional designs on learning. One hundred fourteen college students participated in study 1 and 190 participated in study 2. Students were randomly assigned to one of two conditions in study 1 and one of three conditions in study 2.

2.1. Study 1

The objective of study 1 was to determine how individual differences (background knowledge, verbal ability, visual memory) and instructional designs using narration (*narration* + *image* vs. *narration-only*) would affect students' learning.

2.1.1. Participants

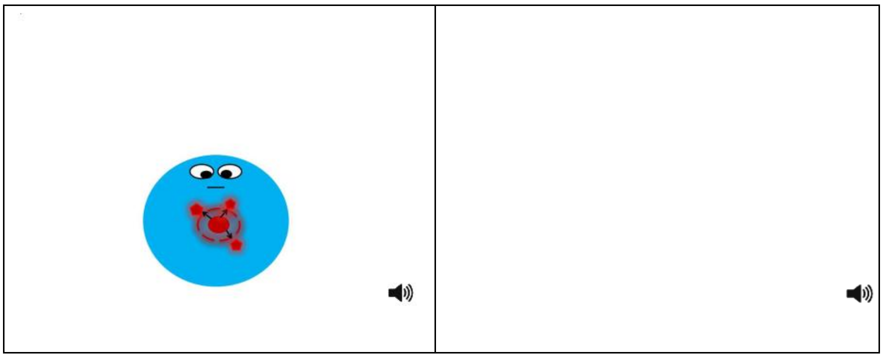
A total 114 college students participated in this experiment. Five students were excluded due to incomplete data. Students were randomly assigned to either the *narration* + *image* group (n = 56) or the *narration-only* group (n = 53). Using a p value of .05, the two groups were not different in age, education, or concept formation ability on the Woodcock Johnson III Test of Cognitive abilities (WJ-III-COG; Woodcock, McGrew, & Mather, 2002).

2.1.2. Assessments

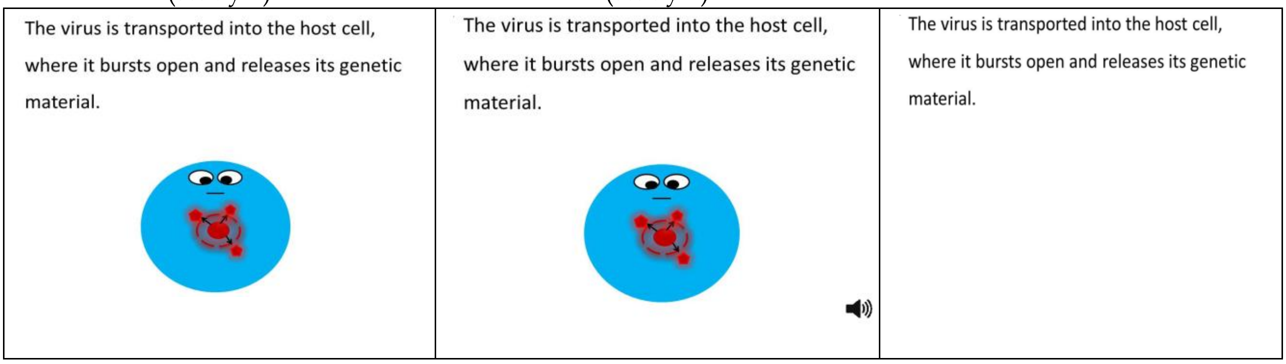
All participants completed three tasks: experimental learning task, verbal ability task, and visual sequential memory task. A description of these measures follows.

Experimental learning task

Learning material and comprehension questions were adapted from Mayer and Estrella (2014) with the first author's permission. The experiment included a single lesson, comprising ten PowerPoint slides which explained how a virus causes a cold. In the *narration* + *image* condition, each slide included an image and an audio clip describing the sequence of events in which a virus enters a body and causes a cold (see Fig. 1 for presentation samples). The *narration-only* condition included the narration in the absence of pictures. After viewing the slides, five open-ended questions were given to students to assess their recall and transfer of information. The retention question was to measure how much learners remember and the transfer question to measure how learners use what they have learned in new problems. The maximum scores were 14 for the retention task and 8 for the transfer task. Related to the experiment, an assessment measure of background knowledge of biology, adapted from Mayer and Estrella (2014), was given to all students. Students self-rated their knowledge of biology on a 5-point scale (1 = very low, 5 = very high) (maximum score = 12).



(a) narration + image condition (study 1)
(b) narration-only condition (study 1)



(c) text + image condition (study 2)
(d) text + image + narration condition (study 2)
(e) text-only condition (study 2)

Figure 1: Examples of slides from each condition in studies 1 and 2
Note: Speaker icon in the *narration + image*, *narration-only*, and *text + image + narration* conditions was not presented in the experiment.

Verbal ability and visual sequential memory

Students' oral language and word knowledge were measured using the verbal ability test from WJ-III-COG (Woodcock et al., 2002). Picture vocabulary, synonyms, antonyms, and verbal analogies subtest scores composed the verbal ability composite. Test-retest reliability was reported at .92 (McGrew, Schrank, & Woodcock, 2007). The visual sequential memory subtest from the *Test of Memory and Learning* (Reynolds & Voress, 2007) was to measure nonverbal memory, learning, and delayed recall. Test-retest reliability was reported at .83 for children and adolescents and .93 for adults (Schmitt & Decker, 2008).

2.1.3. Procedure

Tasks were conducted in the following order: background knowledge questionnaire, verbal ability test, experimental task, and visual sequential memory test. The procedure took 45-60 minutes to complete. For the experimental task, the experimenter presented oral instructions to the students explaining that they would be given an explanation of how a virus attacks a body on a computer screen. Students wore headphones and were seated in a chair in front of a computer. The audio configuration was tested prior to beginning the individual experiment and students were allowed to adjust the audio volume. There was no time limit to answer the questions.

2.1.4. Results

A summary of descriptive data is presented in Table 1. To explore the unique effect of the independent variables on comprehension, hierarchical multiple regression analyses were performed. To control for the effect of background knowledge, background knowledge was entered at step 1. Verbal ability score and visual sequential memory score were entered at step 2 and presentation type (*narration + image* vs. *narration-only*) was entered at step 3. In the absence of statistically significant interaction, we did not include the interaction terms in the model. Comprehension measures (retention and transfer score) served as outcome variables.

	Study 1		Study 2		
	Narration+ image (*N*=56)	Narration-only (*N*=53)	Text+ image (*N*=60)	Text+image+ narration (*N*=63)	Text-only (*N*=60)
Background knowledge (maximum score = 12)	6.14 (2.26)	5.98 (2.31)	4.87 (2.03)	4.63 (1.61)	5.47 (2.20)
Verbal ability (SS, average = 100)	94.59 (8.49)	92.34 (12.17)	93.40 (11.60)	96.15 (10.07)	92.97 (10.31)
Visual sequential memory (SS, average = 10)	10.02 (2.81)	9.70 (2.59)	10.71 (2.68)	10.65 (2.13)	10.45 (2.62)
Retention task (maximum score = 14)	7.59 (2.27)	5.98 (2.24)	7.32 (1.94)	6.22 (2.40)	6.18 (2.03)
Transfer task (maximum score = 8)	2.27 (1.39)	1.77 (1.18)	1.97 (1.25)	1.77 (1.30)	1.35 (1.13)

Table 1: Mean and standard deviation of test results across learning conditions

Note: Standard deviation in parentheses; SS = Standard score; verbal ability score from Woodcock Johnson III Test of Cognitive Abilities; visual sequential score from Test of Memory and Learning, second edition.

Multicollinearity and independent error assumptions were examined prior to the regression analyses. All values of the variance inflation factor (VIF) were closed to 1.0 (less than 2.0), suggesting multicollinearity was not biasing the regression model (Myers, 1990). Values of the Durbin-Watson test were between 1 and 3, suggesting the residuals in the models were independent (Durbin & Watson, 1951). The case wise diagnostic tests indicated that any case did not have a standardized residual larger than 3, indicating the model was not biased by a few outliers. Examination of residual values revealed no issues with normality.

Table 2: Hierarchical regression models predicting comprehension for study 1

Retention task		R^2	*F*	*B*	*SE*	β	*t*
		.22	7.24***				
	Background knowledge			.15	.09	.14	1.57
	Verbal ability			.05	.02	.23	2.45*
	Visual sequential memory			.09	.08	.10	1.09
	Presentation type (*Narration+image* vs. *Narration-only*)			1.43	.41	.30	3.42***
Transfer task		R^2	*F*	*B*	*SE*	β	*t*
		.15	4.68**				
	Background knowledge			.11	.06	.19	2.01*
	Verbal ability			.03	.01	.20	2.09*
	Visual sequential memory			.06	.05	.12	1.31
	Presentation type (*Narration+image* vs. *Narration-only*)			.39	.24	.15	1.64

Note. $N = 109$, * $p < .05$, ** $p < .01$, *** $p < .001$.

For retention task, the model with four variables produced $R^2 = .22$, $F(4, 104) = 7.24$, $p < .001$ (see Table 2). The verbal ability score, $\beta = .23$, $t(104) = 2.45$, $p = .01$, and the presentation type, $\beta = .30$, $t(104) = 3.42$, $p < .001$, significantly predicted the retention task score.

The results indicated that students with higher verbal ability had a higher retention score and the students in the *narration + image* condition had a significantly higher score in the retention task than the students in the *narration-only* condition. Background knowledge or visual sequential memory did not significantly predict the retention score, *p*s > .05. For transfer task, the model produced $R^2 = .15$, $F(4, 104) = 4.68$, $p = .001$. Background knowledge, $\beta = .19$, $t(104) = 2.01$, $p = .04$, and verbal ability score, $\beta = .20$, $t(104) = 2.09$, $p = .03$, significantly predicted the transfer task scores. Students with higher background knowledge and those with higher verbal ability had a significantly higher transfer scores. Visual sequential memory score or presentation type was not significantly related to the transfer score, *p*s > .05.

2.2. Study 2

The aim of study 2 was to determine how the individual differences and instructional design affects learning performance of college students in text-based lessons (*text* + *image*, *text* + *image* + *narration*, *text-only*).

2.2.1. Participants

190 college students participated in this experiment. Because of incomplete testing or technical problems, data from seven students were excluded. Students were randomly assigned to three conditions: *text* + *image* (n = 60), *text* + *image* + *narration* (n = 63), or *text-only* (n = 60). The three groups did not differ in age, education, or concept formation ability, *p*s > .05.

2.2.2. Assessments

Experimental learning task and verbal and visual sequential memory tasks

The learning material and comprehension questions were identical to those used in study 1 (see Fig. 1 for presentation samples). In the *text* + *image* condition, written text along with images were presented. In the *text* + *image* + *narration* condition, text and narration have identical words. In the *text-only* condition, text was presented in the absence of pictures.

2.2.3. Instruments

In study 2, we measured students' eye gaze while they viewed the experimental learning slides. LC Technologies EyeFollower binocular system (sampling rate: 120 Hz, gaze-point tracking accuracy: 0.45 degree) was used to collect eye data. Minimum fixation duration was 100 ms and a spatial dispersion threshold was 1.5°(minimum deviation of 25 screen pixels). Text was presented on a 24 inch computer screen (1,920×1,080 pixels resolution).

2.2.4. Procedures

As in study 1, the experiment was conducted in the following order: background knowledge questionnaire, verbal ability test, experimental task, and visual sequential memory test. After the verbal ability test, the experimenter read instruction to each student describing the eye-tracking procedure, followed by a nine-point calibration accuracy test. There was no time limit for answering these questions.

2.2.5. Results

We performed two separate hierarchical regression analyses, with retention and transfer scores as dependent variables (see Table 1 for descriptive data). To control for the impact of background knowledge, it was entered at step 1. Verbal ability and visual sequential memory scores were entered at step 2 and presentation type was entered at step 3. Because presentation type has three categories (*text* + *image*; *text* + *image* + *narration*; *text-only*), we created dummy variables. *Text-only* group was a baseline group. For the first dummy variable, *text* + *image* condition was coded 1 and others coded 0. For the second dummy variable, *text + image + narration* condition was coded 1 and others 0. In the absence of statistically significant interaction, we did not include the interaction terms in the model. As in study 1, we examined the assumptions prior to the regression analysis. The VIF, Durbin-Watson test, case wise diagnostic test, and residual values revealed no issue with normality, multicollinearity, or independence of error.

Table 3: Hierarchical regression models predicting comprehension for study 2

Retention task		R^2	*F*	*B*	*SE*	β	*t*
		.18	7.61***				
	Background knowledge			.12	.08	.11	1.47
	Verbal ability			.04	.02	.20	2.75**
	Visual sequential memory			.18	.06	.21	2.99**
	Presentation type (*text+image* vs. *text-only*)			1.14	.37	.25	3.10**
	Presentation type (*text+image+narration* vs. *text-only*)			–.02	.38	–.04	–.04
Transfer task		R^2	*F*	*B*	*SE*	β	*t*
		.14	5.60***				
	Background knowledge			.05	.05	.08	1.12
	Verbal ability			.03	.01	.22	2.98**
	Visual sequential memory			.07	.04	.13	1.89
	Presentation type (*text+image* vs. *text-only*)			.63	.22	.24	2.90**
	Presentation type (*text+image+narration* vs. *text-only*)			.36	.22	.14	1.63

Note. N = 183, * $p < .05$, ** $p < .01$, *** $p < .001$.

For the retention task, the equation with five variables (background knowledge, verbal ability, visual sequential memory, presentation type dummy 1, presentation type dummy 2) was significant, $R^2 = .18$, $F(5, 182) = 7.61$, $p < .001$ (see Table 3). Verbal ability score and visual sequential memory score significantly contributed to retention score, β = .20, $t(182) = 2.75$, $p = .007$ for verbal ability and β = .21, $t(182) = 2.99$, $p = .003$ for visual sequential memory. Dummy 1 was significant, β = .25, $t(182) = 3.10$, $p = .002$, indicating that students in the *text + image* condition had significantly higher retention scores than students in the other conditions. Background knowledge or dummy 2 was not significantly related to retention score, *p*s > .05. For the transfer task, the inclusion of the five variables accounted for significant variance in the score, $R^2 = .14$, $F(5, 182) = 5.60$, $p < .001$. Verbal ability score and presentation type dummy 1 were the significant predictor, β = .22, $t(182) = 2.98$, $p = .002$ for verbal ability and β = .24, $t(182) = 2.90$, $p = .004$ for presentation type dummy 1.

Background knowledge, visual sequential memory score, or dummy 2 was not significant, *p*s > .05. To compare performances of the *text* + *image* and the *text* + *image* + *narration* groups, we also conducted the regression using contrast coding (Overall & Spiegel, 1969). For the retention task, students in the *text* + *image* condition had significantly higher score than students in the *text* + *image* + *narration* condition, $t(177) = 3.20$, $p = .002$. For the transfer task, even though students in the *text* + *image* condition had higher scores than the students in the *text* + *image* + *narration* condition, the difference did not reach a significant level, $t(177) = 1.88$, $p > .05$.

Second, we compared eye fixation times for two regions (text and image) on the presentation screen corresponding to the presentation type (*text* + *image* vs. *text* + *image* + *narration*). We did not include the *text-only* condition in the analysis because it did not have the image region. A univariate analysis of variance (ANOVA) was performed with presentation type as an independent variable and total gaze duration as a dependent variable. Total gaze duration was defined as the sum of the durations of fixations. On average, students in the *text* + *image* group spent 1956.64 ms (*SD* = 901.31) on an image region and 4605.45 ms (1107.85) on a text region. Students in the *text* + *image* + *narration* group spent 2036.33 ms (726.14) on an image region and 4187.32 ms (1051.1) on a text region. That is, in the *text* + *image* condition, students spent 30% of time in the image and 70% in the text. In the *text* + *image* + *narration* condition, students spent 33% of time in the image and 67% in the text. The ANOVA yielded only a marginally significant group difference in the text region, $F(1, 118) = 4.50$, $p = .04$ and no difference in the image region, $F(1, 118) = 0.28$, $p > .05$.

3. Discussion

The purpose of our study was to investigate college students' learning with respect to learner characteristics and instructional design by analyzing students' comprehension and eye gaze patterns. All students were assessed in the areas of background knowledge, verbal abilities, and visual sequential memory. In study 1, instructional lessons were narration-based presentations (*narration* + *image* and *narration-only*) while in study 2, lessons were text-based presentations (*text* + *image*, *text* + *image* + *narration*, and *text-only*). In both experiments, students' verbal abilities consistently predicted learning outcomes across the presentation conditions. In addition, both behavioral and eye gaze pattern data supported that instructional design affected students' learning. In contrast with previously reported studies, students' visual sequential memory and background knowledge had little impact on their learning outcomes.

Theoretical and research implications

Our finding for the role of verbal ability in learning is consistent with Gernsbacher's (1997, 2016; Gernsbacher, Robertson, Palladino, & Werner, 2004) structural building framework which posits that general verbal ability regardless of learning modality provides the basis for building coherent mental structures. Scheiter, Schüler, Gerjets, Huk, and Hesse (2014) studied high school students' comprehension during biology instruction in either a text-only or a text + animation condition. Irrespective of the type of media used, students' language comprehension skills strongly predicted their learning outcomes. Zacks, Speer, and Reynolds (2009) found that reading comprehension and film comprehension were based on the same event structure mechanisms. Taken together, these previous works in conjunction with evidence from the current studies suggest that multimedia comprehension is at least partially dependent on the strengths of learner's verbal abilities.

Another possible explanation for the link between verbal ability and learning outcomes is that textual representation is often used to convey more specific information than other instructional forms. In the current study, pictures provided general depictions of important concepts while text described the key concepts with greater specificity. In the acquisition of scientific knowledge, text is the most frequently used medium for instruction whether it stands alone or serves as an adjunct to pictures (Van den Broek, 2010). In a recent paper, Cohn, Taylor, and Pederson (2017) categorized interactions between the verbal and the visual modalities into three types: (a) autonomous, (b) dominant, and (c) assertive. Autonomous types are unimodal. The verb-autonomous type occurs when text is presented in the absence of pictures, while the vis-autonomous type is represented by a sequence of pictures in the absence of text. The dominant interaction is multimodal with one modality serving the dominant role. For example, a verb-dominant interaction is represented by a semantically dominant text supplemented by pictures that are largely redundant or decorative. Finally, the assertive interaction type involves multiple modalities, each of which is necessary for conveying meaning. Further research is needed to replicate this taxonomy and to incorporate a wider range of instructional designs with learning materials that are less dependent on the verbal modality (e.g., visual-dominant or assertive types).

Mayer's principles for multimedia learning and redundancy were supported in the current studies. After controlling for scores in verbal ability and background knowledge, presentation type still plays a significant and unique role in students' learning. Students' performance was better in the *narration + image* condition than in the *narration-only* condition and better in the *text + image* condition than in the *text-only* condition, in line with Mayer's (2002) multimedia principle. Also, in line with Mayer's (2002) redundancy principle, redundant presentation of information (*text + image + narration*) did not improve students' learning. Finally, our findings regarding eye gaze patterns revealed that there was little difference in the time spent viewing the text or image areas when the *text + image* and the *text + image + narration* groups' conditions were compared. Even though students in the *text + image + narration* group were able to listen to the narration, leaving them greater time to process the pictures, they did not take advantage of all additional resources. Instead, they spent time viewing the text while listening (e.g., splitting their attention between two redundant sources), which might have negatively affected their learning. Yue, Bjork, and Bjork (2013) suggest that redundant presentation of text and narration promotes surface-level processing (i.e., word-to-word comparisons), diminishing the use of mental resources for meaningful comprehension.

In both experiments, background knowledge and visual memory had little impact on the students' learning. Coiro (2011) noted that the impact of topic-specific background knowledge on learning diminishes when learning strategies become more important. Willson and Rupley (1996; 1997) found that at the elementary grades, content or topic knowledge had a high impact on reading comprehension but that the impact decreased as use of reading strategies became increasingly important as students advanced in their grade levels. In our two studies, students possessed overall low background knowledge, precluding our ability to fully explore the degree to which knowledge status might differentially affect the learning. The students' low background knowledge could explain low scores on the transfer task. Kalyuga (2005) found that students with little background knowledge were less capable of creating more abstract mental models possibly resulting in their weakened ability to answer questions on transfer tasks that require the application of knowledge learned to new situations or problems.

The fact that visual sequential memory did not affect learning outcomes in this study may be a consequence of the verbal-dominant interaction type (described above) of learning material and tasks. Had we included a task that required the acquisition of perceptual knowledge (e.g., recalling of visual features or understanding movement patterns), visual memory would likely have contributed to learning outcomes as did verbal ability in the current experiments. In fact, in a study of visual abilities and learning outcomes, Imhof, Scheiter, and Gerjets (2011) presented college students with four different locomotion patterns of fish, followed by a pictorial locomotion pattern classification test. In the test, students viewed pictures of fish and determined their locomotion patterns using what they learned in the learning phase. Students with high visual abilities performed better in the locomotion pattern classification tests when compared to students with weaker visual abilities. This study along with the previous studies supports that learner characteristics interact with learning material or task types, differentially affecting learning outcomes.

Practical implications

Our findings suggest that students with low verbal ability may have difficulty with multimodal learning. Indeed, previous studies (e.g., Kim, Wiseheart, & Walden, 2018; Parmar & Signer, 2005) have shown that students with reading deficits are less proficient in processing multimedia-delivered information. In addition, pictures can be seductive and divert learners' attention away from key information (Brunyé, Taylor, Rapp, & Spiro, 2006; Tsai, Wu, & Chen, 2019). Recent research on instructional strategies offers recommendations on how to improve multimedia learning, but most recommendations focus on facilitating students' visual scanning of displays. Here, we will introduce a few recommendations with respect to learner's verbal abilities.

First, students need to have numerous opportunities to learn and exploit general academic words and discipline-specific words since students' language proficiency is strongly linked to overall achievement including mathematics (Vukovic & Lesaux, 2013) and science (Nagy & Townsend, 2012; Tong et al., 2014). Content-specific academic words and across-discipline academic words should be taught in meaningful contexts such as reading expository texts or engaging in oral (e.g., classroom debates) and written (e.g., journalism reports) language activities.

Second, students should be encouraged to spend adequate time on specific target areas (e.g., text or pictures) while processing multimedia information. In a study by Stalbovs, Scheiter, and Gerjets (2015), a group of college students were asked to use an "if-then" implementation intention (e.g., '*If I have finished reading, I will find related information in the image' or 'If I have looked at an image, I will find the text that explains what the image described.')* directly before learning new information. A control group was not taught the implementation intention. The experimental group using the implementation intention strategy had significantly higher scores on tests of recall and reasoning than the control group.

Limitations and direction for future research

First, visual ability, an independent variable in the current study, can be assessed in different ways. We measured students' ability to recall and order visual information in a sequential manner because our participants viewed procedural information through multimedia presentations. However, as noted by Höffler (2010), there are multiple subareas in the visual-spatial domain that include spatial visualization, spatial relations, closure speed, flexibility of closure, and perceptual speed. Future studies might include a wider range of visual abilities to determine the potentially different roles of varying types of visual ability.

Second, our study targeted a general population of college students to explore relationships between media types, individual characteristics (visual and verbal ability), and learning outcomes. Follow-up studies that include students with low reading skills or short attention span might yield different results. For example, for students with academic challenges, redundant information might help them strengthen their mental models constructed from each segment, thereby reducing cognitive load rather than dividing their attention. In fact, Lin, Lee, Wang, and Lin (2016) recently reported a reverse redundancy in their study of second language learning. When students learning English as a second language were taught scientific concepts in an animated narration condition with subtitles or in an animated narration-only condition, students who received the multimedia instruction performed significantly better. Similarly, Knoop-van Campen, Segers, and Verhoeven (2018) found a reversed redundancy effect for children with dyslexia.

Finally, in our study, the multimedia lesson was relatively short. Future studies should use instructional materials that require a longer time to learn, allowing researchers the opportunity to observe if behaviors such as motivation, attention, and inhibition moderate the efficacy of multimedia learning.

4. Conclusion

Our study underscores the important role of individual cognitive differences when investigating learning while also supporting that instructional design affects new learning. Verbal ability was a strong mediator regardless of instructional modality type, suggesting that learners' verbal ability and learning from narration, text, and multimedia are closely interrelated. This finding implies that, for learners with weak verbal ability, multimedia instruction is not a panacea. With respect to media type, multimedia presentations generated better learning performance than single media presentations, irrespective of whether information was presented through narration or text. Finally, redundant presentations of information did not improve learning in college students. Considering varying formats of instruction should be accompanied by strengthening learners' verbal skills and supporting the formation of mental structures for the content at hand during multimedia instructional lessons.

5. Acknowledgements

We would like to express our great gratitude to members of the EyeRead Research Lab at the University of Central Arkansas: Erin Tucker, Lacey Drennan, Seoyoung Kim, Caylin Thompson, Beatriz Cope, and Chelsi Benton.

References

Brunyé, T. T., Taylor, H. A., & Rapp, D. N. (2008). Repetition and dual coding in procedural multimedia presentations. *Applied Cognitive Psychology*, *22*(7), 877–895. https://doi.org/10.1002/acp.1396

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, *20*(7), 917–940. https://doi.org/10.1002/acp.1236

Chiu, T. K. F., & Churchill, D. (2015). Exploring the characteristics of an optimal design of digital materials for concept learning in mathematics: Multimedia learning and variation theory. *Computers & Education*, *82*, 280–291. https://doi.org/10.1016/j.compedu.2014.12.001

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/bf01320076

Clark, R. C., & Mayer, R. E. (2016). *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning* (4th ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.

Cohn, N., Taylor, R., & Pederson, K. (2017). A picture is worth more words over time: Multimodality and narrative structure across eight decades of American superhero comics. *Multimodal Communication*, *6*(1), 19–37. https://doi.org/10.1515/mc-2017-0003

Coiro, J. (2011). Predicting reading comprehension on the internet: Contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research*, *43*(4), 352–392. https://doi.org/10.1177/1086296X11421979

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika*, *30*(3/4), 159–178. https://doi.org/10.2307/2332325

Gary, L., Thomas, N., & Lewis, L. (2010). *Teachers' Use of Educational Technology in U.S. Public Schools: 2009 (NCES 2010-040)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Gernsbacher, M. A. (1997). Two Decades of Structure Building. *Discourse Processes*, *23*(3), 265–304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A. (2016). *Language comprehension as structure building*. New York, NY: Routledge.

Gernsbacher, M. A., Robertson, R. R. W., Palladino, P., & Werner, N. K. (2004). Managing Mental Representations During Narrative Comprehension. *Discourse Processes*, *37*(2), 145–164. https://doi.org/10.1207/s15326950dp3702_4

Goetz, E. T., Sadoski, M., Stricker, A. G., White, T. S., & Wang, Z. (2007). The role of imagery in the production of written definitions. *Reading Psychology*, *28*(3), 241–256. https://doi.org/10.1080/02702710601186381

Goolsby, R. D., & Sadoski, M. (2013). A theoretical approach to improving patient education through written materials. *Annals of Behavioral Science and Medical Education*, *1*(19), 14–18. https://doi.org/10.1007/BF03355232

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. *Educational Psychology Review*, *22*(3), 245–269. https://doi.org/10.1007/s10648-010-9126-7

Imhof, B., Scheiter, K., & Gerjets, P. (2011). Learning about locomotion patterns from visualizations: Effects of presentation format and realism. *Computers & Education*, *57*(3), 1961–1970. https://doi.org/10.1016/j.compedu.2011.05.004

Jared, D., Poh, R. P. Y., & Paivio, A. (2013). L1 and L2 picture naming in Mandarin–English bilinguals: A test of bilingual dual coding theory. *Bilingualism: Language and Cognition*, *16*(2), 383–396. https://doi.org/10.1017/S1366728912000685

Jessen, F., Heun, R., Erb, M., Granath, D.-O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, *74*(1), 103–112. https://doi.org/10.1006/brln.2000.2340

Jones, L. C. (2009). Supporting student differences in listening comprehension and vocabulary learning with multimedia annotations. *CALICO Journal; San Marcos*, *26*(2), 267–289. https://doi.org/10.1558/cj.v26i2.267-289

Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (1st ed., pp. 325–338). New York, NY: Cambridge University Press.

Kim, S., Wiseheart, R., & Walden, P. (2018). Do multimedia instructional designs enhance comprehension in college students with dyslexia? *Journal of Postsecondary Education and Disability*, *31*(4), 351–365.

Knoop-van Campen, C., Segers, E., & Verhoeven, L. (2018). The modality and redundancy effects in multimedia learning in children with dyslexia. *Dyslexia*, *24*(2), 140–155.

Lin, J. J. H., Lee, Y. H., Wang, D. Y., & Lin, S. S. J. (2016). Reading subtitles and taking enotes while learning scientific materials in a multimedia environment: Cognitive load perspectives on EFL students. *Educational Technology & Society*, *19*(4), 47–58.

Lin, P. C., & Yang, C. M. (2010). Impact of product pictures and brand names on memory of Chinese metaphorical advertisements. *International Journal of Design*, *4*(1), 57–70.

Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education*, *81*(3), 356–384. https://doi.org/10.1080/00220973.2012.727885

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85–139. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12–18. https://doi.org/10.1016/j.learninstruc.2014.02.004

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical manual, Woodcock-Johnson III Normative Update*. IL: Riverside: Rolling Meadows.

Myers, R. (1990). *Classical and modern regression with applications*. Boston, MA: Duxbury.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Overall, J. E., & Spiegel, D. K. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin*, *72*(5), 311–322. https://doi.org/10.1037/h0028109

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.

Paivio, A. (2014). Intelligence, dual coding theory, and the brain. *Intelligence*, *47*, 141–158. https://doi.org/10.1016/j.intell.2014.09.002

Parmar, R. S., & Signer, B. R. (2005). Sources of error in constructing and interpreting graphs: A study of fourth- and fifth-grade students with LD. *Journal of Learning Disabilities*, *38*(3), 250–261. https://doi.org/10.1177/00222194050380030601

Pazzaglia, F., Toso, C., & Cacciamani, S. (2008). The specific involvement of verbal and visuospatial working memory in hypermedia learning. *British Journal of Educational Technology*, *39*(1), 110–124. https://doi.org/10.1111/j.1467-8535.2007.00741.x

Reynolds, C. R., & Voress, J. K. (2007). *Test of Memory and Learning-Second Edition (TOMAL-2)*. PRO-ED.

Rupley, W. H., & Willson, V. L. (1996). Content, domain, and word knowledge: Relationship to comprehension of narrative and expository text. *Reading and Writing*, *8*(5), 419–432. https://doi.org/10.1007/BF00404003

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology*, *85*(2), 291–304. https://doi.org/10.1037/0022-0663.85.2.291

Sadoski, M., & Paivio, A. (2013). *Imagery and Text: A Dual Coding Theory of Reading and Writing* (2nd ed.). New York, NY: Routledge.

Scheiter, K., Schüler, A., Gerjets, P., Huk, T., & Hesse, F. W. (2014). Extending multimedia research: How do prerequisite knowledge and reading comprehension affect learning from text and pictures. *Computers in Human Behavior*, *31*, 73–84. https://doi.org/10.1016/j.chb.2013.09.022

Schmitt, A. J., & Decker, S. L. (2008). Test Reviews: Reynolds, C., & Voress, J. K. (2007). "Test of Memory and Learning: Second Edition." Austin, TX: PRO-ED. *Journal of Psychoeducational Assessment*, *27*(2), 157–166. https://doi.org/10.1177/0734282908322449

Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: Using if-then plans to facilitate cognitive processing. *Learning and Instruction*, *35*, 1–15. https://doi.org/10.1016/j.learninstruc.2014.09.002

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, *85*(4), 698–739. https://doi.org/10.3102/0034654314566989

Tong, F., Irby, B. J., Lara-Alecio, R., Guerrero, C., Fan, Y., & Huerta, M. (2014). A randomized study of a literacy-integrated science intervention for low-socio-economic status middle school students: Findings from first-year implementation. *International Journal of Science Education*, *36*(12), 2083–2109. https://doi.org/10.1080/09500693.2014.883107

Tsai, M.-J., Wu, A.-H., & Chen, Y. (2019). Static and dynamic seductive illustration effects on text-and-graphic learning processes, perceptions, and outcomes: Evidence from eye tracking. *Applied Cognitive Psychology*, *33*(1), 109–123. https://doi.org/10.1002/acp.3514

Van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*, 453–456. https://doi.org/10.1126/science.1182594

Vukovic, R. K., & Lesaux, N. K. (2013). The relationship between linguistic skills and arithmetic knowledge. *Learning and Individual Differences*, *23*, 87–91. https://doi.org/10.1016/j.lindif.2012.10.007

Willson, V. L., & Rupley, W. H. (1997). A structural equation model for reading comprehension based on background, phonemic, and strategy knowledge. *Scientific Studies of Reading*, *1*(1), 45–63. https://doi.org/10.1207/s1532799xssr0101_3

Woodcock, R., McGrew, K., & Mather, N. (2002). *Woodcock-Johnson III tests of cognitive abilities*. Itasca, IL: Riverside Publishing Company.

Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology*, *105*(2), 266–277. https://doi.org/10.1037/a0031971

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology. General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305