

## **Picture Superiority Effect: Using Images for Vocabulary Enhancement Among Year One Malaysian ESL Learners**

**Debbita Ai Lin Tan, Bee Choo Lee\*, Shaidatul Akma Adi Kasuma and Malini Ganapathy**

*School of Languages, Literacies and Translation, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*

### **ABSTRACT**

As depicted in the current Malaysian scenario, competency in English is not assured although students are formally taught the language from their first year of schooling in National schools. Malaysian pupils have been achieving English proficiency level lower than the set expectation. This is further highlighted in the 2016 Economic Planning Unit report that after finishing Year 6 of Malaysian primary education, less than half of the students achieved a reasonable level of English literacy. Given the significance of vocabulary knowledge to language proficiency, the present study sought to examine the use of images in Malaysian ESL classrooms in terms of its effects on the development of vocabulary knowledge among Year One pupils'. The study is different from previous studies and exploratory in nature, in the sense that it examined not only receptive vocabulary knowledge but also the productive aspect of vocabulary learning. Two groups of Malaysian Year One pupils from a national primary school in Malaysia were selected by purposive

sampling for a pretest-posttest-delayed posttest design in a quasi-experimental approach. The Experimental group experienced the interventional approach (use of images) while the Control group did not. Results revealed superior productive vocabulary scores, at both the posttest and delayed posttest levels, in favour of the Experimental group. No significant difference was recorded for the receptive dimension. One other finding is that the testing instruments used were statistically

### ARTICLE INFO

#### *Article history:*

Received: 10 December 2019

Accepted: 30 January 2020

Published: 18 March 2020

#### *E-mail addresses:*

debbita\_tan@usm.my (Debbita Ai Lin Tan)

beechoo\_lee@usm.my (Bee Choo Lee)

shaidatul@usm.my (Shaidatul Akma Adi Kasuma)

malinik@usm.my (Malini Ganapathy)

\*Corresponding author

found to be reliable for use with Malaysian Year One pupils. This study and its findings are of relevance to policymakers, educators, curriculum designers, and scholars engaged in ESL/EFL research pertaining to young learners.

*Keywords:* ESL/EFL, images, Picture Superiority Effect, vocabulary, young Malaysian learners

## INTRODUCTION

An average Malaysian student who experiences preschool, primary, secondary and tertiary formal education would have gained approximately an average of 14 to 15 years of English language learning. Ironically, these young Malaysians, in general, have been found to be below the English language competency level expected of them (Economic Planning Unit [EPU], 2016). Despite massive efforts by stakeholders and the government, the issue has become increasingly critical to the extent that this is posing a threat to the realisation of the nation's socio-economic goals. The situation is further compounded by globalisation and internationalisation, which have placed a tremendous amount of expectations on the Malaysian education system to measure up to international standards as well as the global demands of the 21st century. Where English is concerned, the onus is on the English Language Standards and Quality Council (ELSQC), not only to uphold and improve the English language competency of our young but also to align local English standards with that of an internationally

recognised one, the CEFR (Common European Framework of Reference for Languages).

Consistent with the goals of the Malaysian Education Blueprint, the proposed intervention of this research seeks to significantly enhance the vocabulary knowledge of Malaysian Year 1 pupil of various ethnicities, so that their English language acquisition can be effectively fulfilled. In order to achieve this, the research will examine if pupils' receptive and productive vocabulary knowledge is significantly enhanced after intervention implementation.

Vocabulary knowledge has been established to be significant in improving language proficiency and is positively correlated to language development (Cunningham & Stanovich, 1997; Laufer, 2005; Lewis, 2000; Nation, 1990; Wilkins, 1972), and the lack of focus on vocabulary teaching in our Malaysian classrooms, the proposed study is deemed a worthwhile endeavour.

## Research Problem

Evidently, as depicted in the current Malaysian scenario, competency in English is not guaranteed despite being included in the Malaysian curriculum from the very start of primary education which is a Year 1 in both national and national-type schools. Hazita (2016) observed that Malaysian pupils' competence in English left much to be desired, and emphasised the mandated importance of the possession of competent English skills, despite its contested position

dating back to our independence. It is not an inadequate investment in the promotion of this language that has resulted in pupils still dropping below the established literacy standards and requiring further remediation and increased enhancement. As evident in the substantial amounts allocated by the Malaysian government to numerous reformations in the teaching of English language in the last three decades, an ELE standards elevation is clearly visible in every new reform. Take, for instance, the curriculum reform targets a minimum of 90% literacy acquisition after about 3 years of lower primary education (end of Year 3) on the LINUS LBI 2.0 programme based on 12 constructs (Ministry of Education [MOE], 2014) of which the twelfth is able to construct sentences with guidance. Yet, pupils still fall below these established language standards.

Furthermore, in Hayes' (2014) assessment of English provision in primary schools, studies by various researchers (Bauldauf et al., 2011; Kaplan et al., 2011; Nunan, 2003) on international comparisons in the Asia Pacific region have revealed that as far as learners at primary levels are concerned, it is futile to lower the age at entry level for English language learning as it fails as an attempt to increase the dire levels of proficiency.

In research of language acquisition specifically of a second and foreign language, findings that prove the importance of vocabulary in language proficiency is becoming more common. There could be differing views on the most effective way

of language teaching and learning, however, it does not diminish the significance of vocabulary knowledge in developing language competence. Schmitt (2008) observed that the importance of vocabulary growth in improving language proficiency was acknowledged by researchers of second and foreign language studies.

Unfortunately, the teaching of English vocabulary is not specified in the MOE's *Dokumen Standard Kurikulum dan Pentaksiran* and while there is a word list that teachers can refer to, there is no specific model or guideline that they can rely on to teach English vocabulary. Due to the frequency of tests and the intense focus on students excelling in examinations in Malaysia, Ali (2003) pointed out the phenomenon of 'teaching to the tests' (Nuttall, 1995; Norris, 1993). This involves teachers frequently resorting to the drilling or rote-learning method (memorisation) which has been suggested to be less useful to students as mere memorisation may not encourage deep information processing that can result in knowledge being better retained (McQuirter-Scott, 2010).

Meanwhile, the Standards-Based English Language Curriculum emphasises the modular approach and covers the four language components of reading, writing, listening and speaking. The aspects of grammar, phonics and language arts are also accorded focus - but again, none for vocabulary. As advanced earlier, Malaysian students' English language proficiency levels are in dire straits and given the essential dependence of language

proficiency on vocabulary, it is only sensible that an effective approach to vocabulary development is identified in the best interest of our students from the commencement of their formal schooling years. As such, the researchers are proposing to study the effectiveness of using images for vocabulary knowledge development among Malaysian Year 1 pupils, as an effort to discover the most effective method to enhance vocabulary knowledge among Malaysian students and subsequently improve their competence in the English language as aspired by the Malaysian education system. Additionally, a comprehensive search of extant literature has revealed that there is no published research to date, within the Malaysian context, implementing this vocabulary knowledge development method. Not only does this research seek to assess receptive and productive vocabulary knowledge, but it also measures vocabulary knowledge recall (short-term memory

storage of vocabulary knowledge gained) and vocabulary knowledge retention (long-term memory storage of vocabulary knowledge gained). Previous research in the domain of vocabulary learning mainly stops at the recall stage and assessing only receptive vocabulary knowledge.

### Research Objectives and Questions

Broadly, this study examines the use of images in Malaysian primary school level ESL classrooms in terms of its effects on pupils' receptive and productive vocabulary knowledge development. The pupils' pretest, posttest and delayed posttest mean scores, standard deviation values, and significance (*p*) values were compared to determine performance levels between the Experimental and Control groups. Performances were measured at the recall level and the retention level. Table 1 below details the study's objectives and questions.

Table 1  
*Objectives and questions*

Item	Research Objectives	Research Questions
1	To examine if Malaysian Year 1 pupils' receptive vocabulary knowledge recall (short-term) is enhanced with the intervention.	Are Malaysian Year 1 pupils' receptive vocabulary knowledge recall significantly enhanced after undergoing the intervention?
2	To examine if Malaysian Year 1 pupils' receptive vocabulary knowledge retention (long-term) is enhanced with the intervention.	Are Malaysian Year 1 pupils' receptive vocabulary knowledge retention significantly enhanced after undergoing the intervention?
3	To examine if Malaysian Year 1 pupils' productive vocabulary knowledge recall (short-term) is enhanced with the intervention.	Are Malaysian Year 1 pupils' productive vocabulary knowledge recall significantly enhanced after undergoing the intervention?

Table 1 (Continued)

Item	Research Objectives	Research Questions
4	To examine if Malaysian Year 1 pupils' productive vocabulary knowledge retention (long-term) is enhanced with the intervention.	Are Malaysian Year 1 pupils' productive vocabulary knowledge retention significantly enhanced after undergoing the intervention?

## Literature Review

**Vocabulary.** The term vocabulary refers to the knowledge of words and word meanings. Vocabulary includes receptive and productive vocabulary; vocabulary knowledge begins from being able to recognise a word and then progress to being able to meaningfully apply the word in the right context (Faerch et al., 1984). According to Beck and McKeown (1991), vocabulary knowledge development is gradual from one degree to another from having no knowledge of a word to recognition, understanding its meaning, being able to relate its concept to other words and ultimately being able to use it correctly. Clearly, this renders vocabulary knowledge as multidimensional as proposed by Nation (1990), words have form, meaning, and association with other words. Based on this, it is evident that researchers regard vocabulary knowledge as having breadth and depth as opposed to simplistic zero-sum knowledge attainment.

With respect to vocabulary breadth or in other words receptive vocabulary knowledge, Nation (1990) explained that the more central and familiar feature of this aspect was the reflection of word form recognition through the knowledge of its meaning. Meanwhile, Meara (2009) similar to Nation (1990), more recently defined productive vocabulary knowledge, also

known as vocabulary depth, as one's word-association ability. Meara and Wolter (2004) and Meara and Fitzpatrick (2000) explained that vocabulary depth was reflected when a learner was able to link one word with other individual words. For this study's purposes, both vocabulary forms are covered and are operationalised as follows:

1. Receptive vocabulary (breadth):  
Knowledge of word meanings
2. Productive vocabulary (depth):  
Word-association ability

**The Central Importance of Vocabulary Control.** The emphasis on the importance of vocabulary in the development of English language literacy and competence is commonly concluded within the findings of second and foreign language acquisition research. Hunt and Beglar (2005) identified lexicon as key to language comprehension and use. Also lending support to the importance of lexicon in language learning, Singleton (1999) noted that it was the most challenging aspect for L1 and L2 learners, not the area of broad syntactic principles. Even in the opinion of language learners, it is vocabulary acquisition that is considered to be important as a lack thereof, results in poor receptive and productive language use (Nation, 1990).

Eyckmans (2004) observed that second or foreign language learners appreciated the fact that vocabulary knowledge was crucial in the attainment of language proficiency. The growth in vocabulary is clearly beneficial in the development of overall language competence (Nation, 2001; Yuksel & Kavanoz, 2010) but findings from past studies are not limited to only such relationships as various correlation studies have shown that specific language skills are also enhanced and related reciprocally to vocabulary size.

**Picture Superiority Effect.** Pictures are understood universally without the need for translation across languages and because of this nature of pictures, memory stored based on pictorial stimuli is more lasting than the memory of words (Paivio, 1986). When a picture is introduced with either printed or spoken words, recognition memory for pictures is improved (Bower et al., 1975; Madigan, 1983), especially among young learners (Nelson & Kosslyn, 1976). The distinctive power of picture recognition and memory enhancement due to the use of images compared to storage based on mere words as found by Anderson (2009), Brady et al. (2008), Ally and Budson (2007), and Paivio (1991) has provided recognition to the superiority of pictures over words and is termed as 'picture superiority effect' (PSE). Theorists explained that PSE takes place due to the deeper information processing for pictorial stimuli in comparison to mental processing for words.

To further understand the picture superiority effect, it is useful to look at its elaboration in encoding theories and Transfer-Appropriate Processing (TAP) (McBride & Doshier, 2002). The effectiveness of pictures against words and pictures with words have been studied extensively and one of the most prominent hypotheses produced is Paivio's Dual-Coding Hypothesis. The Dual-Coding Theory proposed by Paivio (1971) suggests that pictures are decoded differently which contributes to its superiority over mere words decoding. Paivio (1971) also stated that pictures were coded verbally and in image form rather than just verbally in the case of words stimuli that were single-coded.

This study is designed quite similarly to Paivio's dual-code model consisting of verbal and nonverbal memory codes that are interdependent and involve deeper processing which leads to improved storage of information. The stage of deep processing of information is supported by many researchers for the benefit of recall and retention. Craik and Lockhart (1972) made an important discovery through their seminal work on cognition, learning, and memory. They proposed a Depth of Processing hypothesis which stated that it was the shallowness of an information processing that determined whether the information was stored in the long term memory and not how long it was stored in short term memory. In addition, they also observed that in general, information such

as pictures presented to the learner that was compatible with cognitive structures already previously developed in the learner's brain has a chance of being processed more deeply and stored in short term as well as long term memories.

**Concept Attainment and Inductive Thinking.** The essence of the vocabulary intervention programme lies in the utilisation of inductive thinking combined with presenting pictorial stimuli which display objects and actions familiar to the learner in order to bring about existing words in their vocabulary whether from listening or speaking experiences. This teaching model encourages learners to inquire about the new words, increase the size of their existing sight-reading and writing vocabularies, discover structural

principles and implementing the tools of observation and analysis when studying the four components of language skills - reading, writing, comprehending and composing (Calhoun, 1999). The steps of intervention in the instructional sequence of the model is cyclical all throughout with autonomy given to the teacher in deciding issues such as the duration for each picture as well as the selection of pictures or images used. In making these decisions, the teacher is guided by factors such as the age and language development of the pupils, and the language learning objectives of the teacher. In essence, this intervention incorporates into its model the concept of using pictures as stimuli for language development as well as implements the inductive thinking and concept attainment models of teaching.

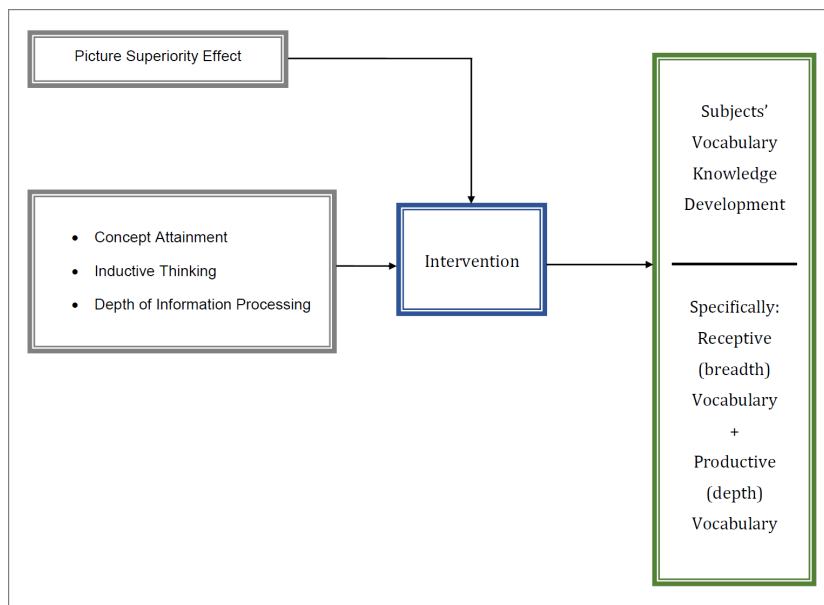


Figure 1. The framework of the study

## METHODOLOGY

### Design

This study focused on Malaysian Year 1 pupil enrolled in Malaysia's national primary schools. The study was specifically designed to examine young learners' receptive and productive vocabulary knowledge development by using images in learning new words. A primarily quantitative approach was adopted for the stated purposes. A quasi-experimental design with purposive sampling formed the core of the study. Quasi-experimental designs are considered worthwhile because they permit researchers to reach reasonable conclusions (Ary et al., 2009) and it is also a design that allows for the collecting evidence of a particular treatment effect especially in the study of impacts in an educational intervention (Diem, 2002). Nested within this design is the pretest-posttest-delayed posttest control-group approach (Revesz et al., 2011). Denzin (1978) observed the importance of 'within-method' triangulation steps such as conducting post-testing and assigning control groups in order to increase research findings credibility.

### Size and Sampling

The participants were male and female Malaysian Year 1 pupil enrolled in a national primary school in Malaysia. The sample size for the study ( $n = 53$ ) is comparable with the range employed by similar past studies on vocabulary knowledge development among children. For example:

Amin Afshar & Mojavezi (2017):  $n = 50$   
Seyed & Azam (2011):  $n = 34$

Sobel et al. (2010):  $n = 39$

Subramaniam et al. (2009):  $n = 20$

According to Gay and Diehl (1992), the recommended number of participants or the ideal sample size for correlational studies or causal-comparative studies which are usually conducted in educational research is 30. They emphasized that the sample size has to be determined based on the types of research being carried out. Causal-comparative studies include studies that examine the effects of interventions (Brewer & Kuhn, 2010). It is important for studies that employ testing designs where multiple groups are involved to have an at least 25 participants so that it is of a good size to observe whether the intervention implemented really has taken effect (Hogg & Tanis, 2005).

The intact classes assigned to the researchers were largely determined by the participating school's administrative authorities, as is common in educational research processes (Ary et al., 2009). However, to adhere as closely as possible to research protocols, the researchers applied randomisation *where possible* and the coin-toss method (Neuman, 2000) was employed to determine the status of each intact class (the Experimental status of Control status).

According to Salkind (2010), intact groups are usually chosen in educational research involving classroom settings. Intact group selection refers to choosing participants based on existing cohorts (such as individuals already organised into classes) (Martella et al., 2013). Intact groups



include language classes, medical or clinical groupings, communities, and organisations.

### Instruments and Intervention

According to vocabulary testing expert, Read (2007), the Yes-No format is the simplest test format and is practical for use. He elaborated that Meara has taken the lead in developing the Yes-No format and making it available for use. The test has been validated by various studies with returned values indicating a positive correlation with other similar tests (Mochida & Harrington, 2006). Meanwhile, the Lex30 Test by Meara and Fitzpatrick (2000) has been repeatedly validated as an adequate instrument to measure ESL learners' productive vocabulary across all ages, including young learners (Gonzalez & Piriz, 2016). Similarly, Fitzpatrick and Clenton (2010) highlighted that the test produces reliable scores, reflects an improvement in language knowledge and produces scores which are comparable to those of similar

tests. The details of the instruments are as presented in Table 2.

Each cycle of the intervention is highly comprehensive as listed out below, based on Calhoun (1999):

1. The teacher puts up the selected picture on the board and gets pupils to sit in groups.
2. The teacher goes through the steps of the intervention with the pupils:
  - Points at an item in the picture and asks pupils to identify what they see;
  - Label by drawing a line from the identified object or area;
  - Pupils say the word and one will write the word on the board;
  - Pupils spell and read the word aloud;
  - When the teacher has gone through the items in the picture, pupils review the picture aloud (read and spell);

Table 2

#### Instruments

Instrument	Type/Source
Image	Selected from Year 1 textbook
Word List	Year 1 Word List provided in MOE's <i>Dokumen Standard Kurikulum dan Pentaksiran</i> (2018)
Tests	Receptive vocabulary: Yes-No Test by Meara (1992) Productive vocabulary: Lex30 Test by Meara and Fitzpatrick (2000)

- The teacher leads pupils into creating a title for the picture;
  - The teacher guides pupils to generate simple sentences using the picture;
  - Reading and reviewing the sentences.
  - The teacher will only induce a few words (enough to generate simple sentences) as an example for the pupils. The pupils will then continue in groups.
3. Pupils experience intervention in their groups.
  4. Teacher gives a copy of the picture to each group with a time limit.
  5. The teacher walks around to guide, supervise, and extend help when necessary (teacher monitors if students get restless and may give breaks if necessary). The teacher also ensures that every pupil gets a chance to participate and that no one is isolated.

### Data Analysis

The administered pre-, post- and delayed post-tests were scored by a qualified teacher. Both descriptive and inferential statistics were employed to lend greater depth to the findings. The Statistical Package for the Social Sciences (SPSS) software (Version 22) was utilised to analyse data gathered from the administration of the tests. Specifically, this software was used to derive the central tendency values or mean scores as well as for standard deviation values (variance of distribution in relation to the mean). SPSS was also employed to run t-tests to check for statistical significance.

The t-test is a comparison between the values of two means to determine whether variables are dependent or not dependent on each other which leads to either accepting or rejecting null hypothesis based on a pre-determined significance level (Agresti & Finlay, 2008). A  $p$ -value that is smaller than a pre-determined threshold value (commonly 0.05) indicates that the observed effect is highly unlikely to have occurred by chance (Privitera, 2012). It is also a measure of the probability that is generalizable to the entire population from which the sample has been derived Dane (2003). The  $p$ -value for the present study is specified at  $p < 0.05$ , a level of significance often employed in psychological and educational studies (Best, 1977).

### Integrity

The researchers, in cooperation with the school's authorities, closely monitored the research to ensure the smooth-running of study protocols. Consent was also obtained from the school's authorities as well as the participants' guardians, and anonymity is preserved to fully protect the participants' privacy.

### Primary Phases

The primary phases of this study are listed in Table 3.

## RESULTS

### Receptive Vocabulary Knowledge (Breadth): Recall and Retention

Both descriptive and inferential statistics

Table 3

*Primary phases*

Phase	Activity	Timeline/Duration
1	Pretesting	Week 1
2	Intervention	Over the course of 2 weeks: 1 session per week, 30 minutes per session ( <i>as permitted by the school's authorities</i> )
3	Post testing	Immediately after the intervention ends
4	Delayed Posttesting	2 weeks after post-testing

were applied to lend greater depth to the findings. The  $p$ -value is set at  $p < 0.05$ .

*RQ 1:* Are Malaysian Year 1 pupils' receptive vocabulary knowledge recall significantly enhanced after undergoing the intervention?

$H_1$ : There is no significant difference in the receptive vocabulary knowledge recall between the Experimental group and the Control group.

*RQ 2:* Are Malaysian Year 1 pupils' receptive vocabulary knowledge retention significantly enhanced after undergoing the intervention?

$H_2$ : There is no significant difference in the receptive vocabulary knowledge retention between the Experimental group and the Control group.

As shown in Table 4, the mean scores of the pretest for both the Control and Experimental groups did not differ significantly, thus denoting baseline similarity at 6.81 ( $SD = 0.392$ ) and 6.74 ( $SD = 0.360$ ) respectively. Posttest data indicate that the Experimental group experienced higher vocabulary breadth recall than the Control group; the Experimental group's mean score at 8.93 ( $SD = 0.388$ ) is superior to the Control group's achievement of 8.77 ( $SD = 0.409$ ). As for the delayed posttest, which measures vocabulary breadth retention, the mean scores achieved were 8.33 ( $SD = 0.453$ ) (Experimental group) and 8.12 ( $SD = 0.478$ ) (Control group), indicating minimal knowledge decay.

At the pretest level, Table 5 shows that the Sig. ( $p$ ) value obtained was more than 0.05 ( $p > 0.05$ ) at  $p = 0.900$ , indicating no

Table 4

*Pretest, posttest and delayed posttest*

	Group	N	Mean	Std. Dev.
Pretest	Control	26	6.81	0.392
	Experiment	27	6.74	0.360
Posttest	Control	26	8.77	0.409
	Experiment	27	8.93	0.388
Delayed Posttest	Control	26	8.12	0.478
	Experiment	27	8.33	0.453

Table 5  
*Independent samples test: Equality of variances and statistical significance*

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Error Diff.	95% Confidence Interval of the Difference	
								Lower	Upper
Pretest (Breadth)	0.006	0.940	0.126	51	0.900	0.067	0.532	-1.001	1.135
Posttest (Breadth)	0.031	0.862	-0.278	51	0.782	-0.157	0.563	-1.288	0.975
Delayed Posttest (Breadth)	0.014	0.908	-0.331	51	0.742	-0.218	0.658	-1.539	1.103

statistically significant difference between group means. Table 5 also indicates that the Sig. ( $p$ ) the value obtained was more than 0.05 ( $p > 0.05$ ) at  $p = 0.782$  at the posttest level, indicating no statistically significant difference between group means at the recall level for vocabulary breadth. Similarly, no statistically significant difference between group means was found at the delayed posttest level (retention) with a value of  $p = 0.742$ . Both  $H_1$  and  $H_2$  are therefore accepted.

The Sig. ( $p$ ) value obtained at the pretest level, which specifies no statistically significant difference between group means, indicates that the testing instrument is reliable in terms of 'reliability as stability over similar samples' (Cohen et al., 2007), thus presenting it as a feasible option for use among similar cohorts in future studies.

### **Productive Vocabulary Knowledge (Depth): Recall and Retention**

Both descriptive and inferential statistics are applied to lend greater depth to the findings. The  $p$ -value is set at  $p < 0.05$ .

*RQ 3:* Are Malaysian Year 1 pupils' productive vocabulary knowledge recall significantly enhanced after undergoing the intervention?

$H_3$ : There is no significant difference in the productive vocabulary knowledge recall between the Experimental group and the Control group.

*RQ 4:* Are Malaysian Year 1 pupils' productive vocabulary knowledge retention

significantly enhanced after undergoing the intervention?

$H_4$ : There is no significant difference in the productive vocabulary knowledge retention between the Experimental group and the Control group.

As demonstrated in Table 6, the mean scores of the pretest for both the Control and Experimental groups did not differ significantly, thus denoting baseline similarity at 2.19 ( $SD = 0.184$ ) and 2.15 ( $SD = 0.183$ ) respectively. Posttest data indicate that the Experimental group experienced higher vocabulary depth recall than the Control group; the Experimental group's mean score at 4.63 ( $SD = 0.234$ ) is superior to the Control group's achievement of 3.38 ( $SD = 0.249$ ). As for the delayed posttest, which measures vocabulary depth retention, the mean scores achieved were 3.67 ( $SD = 0.200$ ) (Experimental group) and 2.38 ( $SD = 0.167$ ) (Control group), indicating the presence of knowledge decay but not at a substantial rate.

At the pretest level, Table 7 shows that the Sig. ( $p$ ) value obtained was more than 0.05 ( $p > 0.05$ ) at  $p = 0.865$ , indicating no statistically significant difference between group means. Table 7 also indicates that the Sig. ( $p$ ) the value obtained was less than 0.05 ( $p > 0.05$ ) at  $p = 0.001$  at the posttest level, indicating a statistically significant difference between group means at the recall level for vocabulary depth. Similarly, Sig. ( $p$ ) value obtained was less than 0.05 ( $p > 0.05$ ) at  $p = 0.000$  at the delayed posttest level, indicating a statistically significant

Table 6

*Pretest, posttest and delayed posttest*

	Group	N	Mean	Std. Dev.
Pretest	Control	26	2.19	0.184
	Experiment	27	2.15	0.183
Posttest	Control	26	3.38	0.249
	Experiment	27	4.63	0.234
Delayed Posttest	Control	26	2.38	0.167
	Experiment	27	3.67	0.200

Table 7

*Independent samples test: Equality of variances and statistical significance*

		Levene's Test for Equality of Variances				
		F	Sig.			
Pretest (Depth)	Equal variances assumed	0.000	0.985			
	Equal variances not assumed					
Posttest (Depth)	Equal variances assumed	0.061	0.805			
	Equal variances not assumed					
Delayed Posttest (Depth)	Equal variances assumed	2.389	0.128			
	Equal variances not assumed					
		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)
Pretest (Depth)	Equal variances assumed	0.000	0.985	0.170	51	0.865
	Equal variances not assumed			0.170	50.960	0.865
Posttest (Depth)	Equal variances assumed	0.061	0.805	3.653	51	0.001
	Equal variances not assumed			3.650	50.662	0.001
Delayed Posttest (Depth)	Equal variances assumed	2.389	0.128	4.905	51	0.000
	Equal variances not assumed			4.923	49.777	0.000

Table 7 (Continued)

		t-test for Equality of Means			
		Mean Diff.	Std. Error Diff.	95% Confidence Interval of the Difference	
				Lower	Upper
Pretest (Depth)	Equal variances assumed	0.044	0.259	-0.477	0.565
	Equal variances not assumed	0.044	0.259	-0.476	0.565
Posttest (Depth)	Equal variances assumed	-1.245	0.341	-1.929	-0.561
	Equal variances not assumed	-1.245	0.341	-1.930	-0.560
Delayed Posttest (Depth)	Equal variances assumed	-1.282	0.261	-1.807	-0.757
	Equal variances not assumed	-1.282	0.260	-1.805	-0.759

difference between group means at the retention level for vocabulary depth. Both  $H_3$  and  $H_4$  are therefore rejected.

The Sig. ( $p$ ) value obtained at the pretest level, which specifies no statistically significant difference between group means, indicates that the testing instrument is reliable in terms of 'reliability as stability over similar samples' (Cohen et al., 2007), thus presenting it as a feasible option for use among similar cohorts in future studies.

## DISCUSSION AND CONCLUSION

It was predicted that the participants who experienced vocabulary learning through images would outperform their peers, in terms of both recall and retention, who learnt

via the conventional approach practised in Malaysian ESL classrooms at the primary school level. The results of the Lex30 Test (productive vocabulary test) strongly support this prediction, with the participants of the Experimental group obtaining higher vocabulary recall and retention scores as compared to the Control group. The results were also found to be statistically significant, indicating that the observed effect is highly unlikely to have occurred by chance.

The results of the Yes-No Vocabulary Test (receptive vocabulary test), meanwhile, showed gains but further analysis returned values indicating the absence of statistical significance. This is possibly due to a testing

issue, in that the target words included in the Yes-No Vocabulary Test were likely too easy for the students which may have caused a ceiling effect to occur and consequently, affected the scores. This is a finding worthy of note, and it is recommended that future studies employing the Yes-No Vocabulary Test takes into account the possibility that a ceiling effect may indeed occur and therefore also consider the inclusion of more complex or difficult target words for testing as well as the use of a larger cohort of participants.

Essentially, productive vocabulary knowledge (depth) development was found to be positive when pupils experienced the intervention, in which they were exposed to images and prompted to establish direct links between the images and corresponding English words, thereby gaining access to two instead of merely one retrieval path (Plass et al., 1998). Apart from augmenting the positive influence of the visual-verbal cognitive style of learning (Koc-Januchta et al., 2017), the findings of the present study are also largely in line with those of similar research. For instance, Wahyuni (2016) whose study was conducted within the context of Indonesia, McDonald (2010) whose study was conducted within the context of Canada, and Calhoun et al. (2001) who framed their research within the American context.

Additionally, the study's findings are fundamentally supportive of Paivio's (1991) dual-code theory; Paivio theorised that pictures, when used with words to elicit a verbal code as well as an image code, have

an advantage over single-coded models in terms of information processing and storage. The study's intervention emphasised on two interdependent types of memory codes, one verbal and the other, non-verbal.

The participants' enhanced vocabulary knowledge recall and retention for the product dimension are also likely due to the picture superiority effect (PSE) which is one's memory for pictures is superior to one's memory for words. The findings of the present study are consistent with previous studies' findings (Anderson, 2009; Ally & Budson, 2007; Brady et al., 2008), in that exposure to visuals can lead to better memory or "welding" of information received. Scholars generally suggest that the deeper level of information processing experienced when presented with pictorial stimuli contribute to PSE.

The study's intervention also capitalises on inductive thinking and concept attainment (Calhoun, 1999). The steps of the intervention are designed to not only undertake the development of sight vocabulary but also to classify new lexical items through concept building and to develop productive abilities. Continuously putting forth questions that stimulate classroom discussion as well as encouraging learners to suggest words from their existing vocabulary are key activities of the intervention. Furthermore, each comprehensive cycle of the intervention is structured around an image and the corresponding target words and activities are intended to maximise student engagement as well as create multiple encounters with the target words. The recorded higher mean



scores of the Experimental group attest to the effectiveness of the intervention's design for lexical development; although this is, with respect to the present study, concentrated on the productive dimension, it is likely that the receptive dimension will also be shown to have benefited once the testing issue is resolved.

Fundamentally, the findings of this study offer vital insights into the effectiveness of the intervention for vocabulary knowledge recall and retention among young Malaysian ESL learners. The essentialist nature of the intervention demonstrates that images, when used to elicit words from a learner's existing vocabulary, can be effective classroom tools for vocabulary instruction.

It is important to note that these findings should not be generalised to the entire (or general) population of young Malaysian ESL learners as the samples in the study were restricted to Year 1 Malaysian ESL learners enrolled in a national primary school in Malaysia. Therefore, the same effect or level of effectiveness may not necessarily be reflected when employed with, for instance, preschool learners. It is recommended that future research look into the effects of this intervention on different and larger cohorts of learners, and also examine its effectiveness with regards to improving ESL students' reading, writing, listening and speaking skills, especially within the Malaysian context in which exists a paucity of such research.

It is to be noted that this study also gleaned one important contribution; the

testing instruments were empirically found to be reliable in terms of 'reliability as stability over similar samples' (Cohen et al., 2007), thus presenting them as feasible options for use among similar cohorts in future studies.

## ACKNOWLEDGEMENT

This research was funded and supported by a Universiti Sains Malaysia (USM) Short Term Grant.

## REFERENCES

- Ally, B. A., & Budson, A. E. (2007). The worth of pictures: Using high-density event-related potentials to understand the memorial power of pictures and the dynamics of recognition memory. *NeuroImage*, 35(1), 378-395.
- Agresti, A., & Finlay, B. (2008). *Statistical methods for the social sciences* (4th ed.). Upper Saddle River, New Jersey: Pearson.
- Ali, M. S. (2003). *English language teaching in primary schools: Policy and implementation concerns*. Retrieved January 30, 2019, from <https://kajianberasaskansekolah.files.wordpress.com/2008/04/article70.pdf>
- Amin Afshar, M. & Mojavezi, A. (2017). The effects of aural and visual storytelling on the vocabulary retention of Iranian EFL learners. *English Language Teaching*, 10(4), 92-99.
- Anderson, J. R. (2009). *Cognitive psychology and its implications* (7th ed.). New York: Worth Publishers.
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. K. (2009). *Introduction to research in education* (8th ed.). Belmont, California: Cengage Learning.
- Baldauf, R. B., Kaplan, R. B., Kamwangamalu, N., & Bryant, P. (2011). Success or failure of primary

- second/foreign language programmes in Asia: What do the data tell us? *Current Issues in Language Planning*, 12(2), 309-323.
- Beck, I. L., & McKeown, M. G. (1991). Conditions of vocabulary acquisition. In P. D. Pearson (Ed.), *The handbook of reading research* (pp. 789-814). New York: Longman.
- Best, J. W. (1977). *Research in education* (3rd ed.). Englewood Cliffs: Prentice-Hall.
- Bower, G. H., Karlin, M. B., & Dueck, A. (1975). Comprehension and memory for pictures. *Memory and Cognition*, 3(2), 216-220.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38), 14325-14329.
- Brewer, E. W., & Kuhn, J. (2010). Causal-comparative design. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 125-132). Thousand Oaks, California: Sage Publications.
- Calhoun, E. (1999). *Teaching beginning reading and writing with the Picture Word Inductive Model*. Virginia: Association for Supervision and Curriculum Development (ASCD).
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). New York: Routledge.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671-684.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33(6), 934-945.
- Dane, F. C. (2003). *Research methods*. Pacific Grove, California: Brooks/Cole.
- Denzin, N. K. (1978). *The research act* (2nd ed.). New York: McGraw-Hill.
- Diem, K. G. (2002). Using research methods to evaluate your extension program. *Journal of Extension*, 40(6), 412-415.
- Economic Planning Unit. (2016). *Strategy paper 10: Transforming education system*. Putrajaya, Malaysia: Author.
- Eyckmans, J. (2004). *Measuring receptive vocabulary size*. Utrecht: LOT.
- Faerch, K., Haastруп, K., & Phillipson, R. (1984). *Learner language and language learning*. Clevedon, England: Multilingual Matters.
- Fitzpatrick, T., & Clenton, J. (2010). The challenge of validation: Assessing the performance of a test of productive vocabulary. *Language Testing*, 27(4), 537-554.
- Gay, L. R., & Diehl, P. L. (1992). *Research methods for business and management*. New York: Macmillan.
- Gonzalez, R. A., & Piriz, A. M. P. (2016). Measuring the productive vocabulary of secondary school CLIL students: Is Lex30 a valid test for low-level school learners? *Vigo International Journal of Applied Linguistics*, 13(1), 31-53.
- Hayes, D. (2014). *Factors influencing success in teaching English in state primary schools*. London: British Council.
- Hazita Azman (2016). Implementation and challenges of English language education reform in Malaysian primary schools. 3L: *The Southeast Asian Journal of English Language Studies*, 22(3), 65-78.
- Hogg, R. V., & Tanis, E. A. (2005). *Probability and statistical inference* (7th ed.). Upper Saddle River, New Jersey: Prentice-Hall.
- Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, 17(1), 23-59.

- Kaplan, R. B., Baldauf, R. B., & Kamwangamalu, N. (2011). Why educational language plans sometimes fail. *Current Issues in Language Planning, 12*(2), 105-124.
- Koć-Januchta, M., Höffler, T., Thoma, G. B., Prechtel, H., & Leutner, D. (2017). Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures - An eye-tracking study. *Computers in Human Behavior, 68*(1), 170-179.
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook, 5*(1), 223-250.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.
- Levie, W. H., & Hathaway, S. N. (1998). Picture recognition memory: A review of research and theory. *Journal of Visual Verbal Language, 8*(1), 6-45.
- Lewis, M. (2000). Introduction. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (p. 10). Hove: Language Teaching Publications.
- Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, memory and cognition* (pp. 65-89). Hillsdale, New Jersey: Lawrence Erlbaum.
- Martella, R. C., Nelson, J. R., Morgan, R. L., & Marchand-Martella, N. E. (2013). *Understanding and interpreting educational research*. New York: Guilford.
- McBride, D. M., & Doshier, B. A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition, 11*(3), 423-460.
- McQuirter-Scott, R. (2010). *What works? Research into practice*. Ontario, CA: The Literacy and Numeracy Secretariat.
- Meara, P. (1992). *EFL vocabulary tests*. Swansea: Centre for Applied Language Studies, University College Swansea.
- Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam: John Benjamins.
- Meara, P., & Fitzpatrick, T. (2000). Lex 30: An improved method of assessing productive vocabulary in an L2. *System, 28*(1), 19-30.
- Meara, P., & Wolter, B. (2004). V\_Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup & B. Henriksen (Eds.), *Angles on the English-speaking world 4* (pp. 85-96). Copenhagen, Denmark: Museum Tusulanum Press.
- Ministry of Education Malaysia. (2014). *Malaysian Education Blueprint Annual Report: 2014*. Kuala Lumpur, Malaysia: Author.
- Mochida, A., & Harrington, M. (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing, 23*(1), 73-98.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, Massachusetts: Heinle.
- Nelson, K. E., & Kosslyn, S. M. (1976). Recognition of previously labelled or unlabeled pictures by 5-year-olds and adults. *Journal of Experimental Child Psychology, 21*(1), 40-45.
- Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4th ed.). Needham Heights, Massachusetts: Allyn and Bacon.
- Norris, N. (1993). Evaluation, economics and performance indicators. In J. Elliott (Ed.),

- Reconstructing teacher education* (pp. 31-38). London: The Falmer Press.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, 37(4), 589-613.
- Nuttall, D. (1995). Examinations in education. In R. Murphy & P. Broadfoot (Eds.), *Effective assessment and the improvement of education* (pp. 33-43). London: The Falmer Press.
- Paivio, A. D. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Paivio, A. D. (1986). *Mental representations*. New York: Oxford University Press.
- Paivio, A. D. (1991). *Images in the mind: The evolution of a theory*. New York: Harvester Wheatsheaf.
- Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *Studies in Second Language Acquisition*, 9(2), 202-221.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, 90(1), 25-36.
- Privitera, G. J. (2012). *Statistics for the behavioural sciences*. Thousand Oaks, California: Sage Publications.
- Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies*, 7(2), 105-125.
- Revesz, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and L2 development. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 203-237). Amsterdam/Philadelphia: John Benjamins.
- Salkind, N. J. (2010). *Encyclopedia of research design*. Thousand Oaks, California: Sage Publications.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
- Seyed Jalal Abdolmanafi-Rokni, & Azam Jannati Atae (2011). The effect of background music on vocabulary recall and retention of children learning English as a foreign language. *International Journal of Basic Sciences & Applied Research*, 3(8), 491-495.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2010). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, 25(5), 763-767.
- Subramaniam, V., Hamzah, M. S. G., Dani, N. A., Abd Rahim, N., Nik Affendi, N. R., & Melebek, A. R. D. (2009). Learning vocabulary through image-sketching among primary school children. *International Journal of Diversity in Organisations, Communities and Nations*, 9(2), 35-42.
- Wahyuni, E. F. (2016). *The use of the Picture Word Inductive Model (PWIM) to improve students' vocabulary achievement at MTs Assyafi'iyah Gondang* (Master's thesis), State Islamic Institute, Indonesia.
- Wilkins, D. A. (1972). *Linguistics and language teaching*. London: Edward Arnold.
- Yuksel, H. G., & Kavanoz, S. (2010). Does vocabulary knowledge distinguish among proficiency levels of Turkish university students? *International Journal of Learning*, 17(5), 513-521.