

Investigating the Relationship Between Visual Attention, Story Comprehension, and Vocabulary Skills in Malaysian Prereaders: An Eye-Tracking Study


Hoo Keat Wong, University of Nottingham in Malaysia, Malaysia*

Siew Ming Thang, HELP University, Malaysia

Chee Hao Sue, HELP University, Malaysia

Rosalind A. K. Ahju, HELP University, Malaysia

Fung Lan Loo, Open University, Malaysia

 <https://orcid.org/0000-0003-2188-3968>

ABSTRACT

Based on the cross-channel connections between auditory and pictorial representations, it has been proposed that the presentation of coherent narration along with the picture and text content may enhance children's story comprehension and vocabulary learning. The authors tested 40 four- to five-year-old Malaysian prereaders (17 Malays, 23 Chinese) for story comprehension while observing their eye movements to determine the degree to which the presence of pictures and/or text aids understanding of the narration and influences looking patterns. Both Malay and Chinese prereaders showed no interest in the printed text that was presented alongside the picture on the same page, which is consistent with earlier findings. This suggests that ethnic origins have little influence on how prereaders direct their visual attention to the relevant information for story comprehension. When there was no narration, they fixated longer on the text and less on the image, indicating that a significant amount of mental effort was required to process the words without verbal information. Regardless of stimulus congruency, storytelling performance affected how much children focused on target objects and keywords. More intriguingly, it was found that in Malay prereaders, there was a correlation between story comprehension and vocabulary skills across tasks. Additionally, Malay prereaders who performed well looked at the displayed stimuli longer than Chinese prereaders who performed well, especially when a narrator was presented alongside the visual stimulus. These novel findings are discussed along with their implications for multimedia learning and future research directions.

KEYWORDS

Attention, Children, Eye-Tracking, Multimedia Learning, Story Comprehension, Vocabulary Skills

COGNITIVE PROCESSES IN MULTIMEDIA LEARNING

Multimedia learning involves using visual and auditory stimuli to convey information to learners. In recent years, there has been a growing interest in understanding the cognitive processes involved

DOI: 10.4018/IJCALLT.332878

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

in multimedia learning. Several empirical studies have investigated how the young brain processes information presented in different multimedia formats, such as videos, animations, and interactive simulations. Some of these studies have examined the effects of multimedia on attention, working memory, and comprehension. For example, Moreno and Mayer (2007) investigated the effect of multimedia on attention and comprehension in college students. The researchers found that students who received multimedia instruction showed greater attention and comprehension compared to those who received text-only instruction. Another study by Mayer, Heiser, and Lonn (2001) examined the effect of multimedia on working memory, revealing that multimedia instruction resulted in better performance on tasks that required working memory.

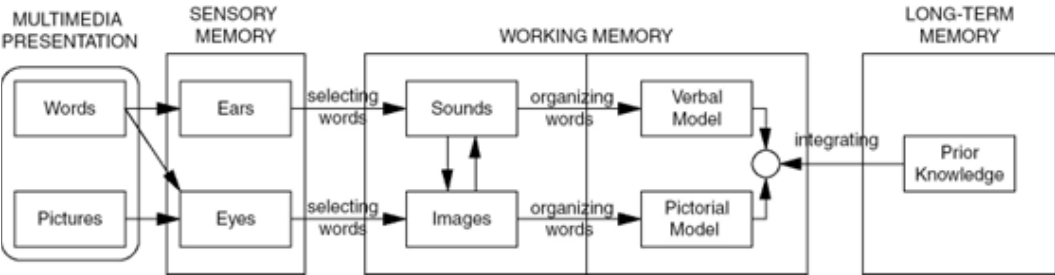
Sharing picture storybooks with young children is crucial for developing literacy skills. Many children's storybooks are illustrated via visual and/or oral narratives. When children read or listen to a story, both visual (pictorial) and auditory (verbal) processing channels are employed. Several cognitive theories and models have been proposed and further adopted by educators and modern researchers in empirical studies to understand the cognitive processes underlying a child's story comprehension and multimedia learning. According to Paivio's (1986) Dual Channel Model, humans possess two channels to process visual and auditory information. The first is the verbal channel, which processes spoken words heard through the ears; the other is the visual channel, which processes images seen through the eyes (Paivio, 2007). Each channel processes and stores information differently, with the visual channel being specialised in processing and storing visual information and the verbal channel being specialised in processing and storing verbal information. The Dual Channel Model suggests that multimedia learning is most effective when it utilises both channels, as this creates a stronger memory trace and facilitates better comprehension of the information. This idea is particularly relevant when considering multimedia learning in young children, as the use of both visual and verbal information can help them better understand and retain the information being presented.

The cognitive theory of multimedia learning (CTML; Mayer, 2001) further explains how children process information when reading or hearing a story via multimedia. That is, the information in both visual and verbal form contributes to the comprehension of the story and increases the learner's visual attention. CTML proposes a model that comprises three memory stores: sensory memory, working memory, and long-term memory (see Figure 1). In working memory, the auditory information is organised to construct a verbal model; meanwhile, the images will be processed to create pictorial coding to build a visual model. When these two processes occur concurrently, integrating the auditory and visual channels will increase the efficiency of the information processing. This effect, in turn, facilitates the comprehension of the learners (Sorden, 2013).

In the CTML, Mayer also proposed several principles, such as the multimedia principle, modality principle, redundancy principle, and coherence principle, to guide the design of multimedia learning

Figure 1. Mayer's (2001) cognitive theory of multimedia learning

Note. Adapted from "Applying the Science of Learning: Evidence-Based Principles for the Design of Multimedia Instruction" by Mayer, R. E., 2008, The American Psychologist, 63(8), p. 762. Copyright 2008 by the American Psychological Association.



materials. These principles aim to optimise the cognitive processing of information and reduce extraneous cognitive load, making them complementary and compatible with the ideas proposed by the Cognitive Load Theory (CLT; Sweller, 1994). CLT serves as a framework to understand and optimise the learning process. It is grounded in the idea that human cognitive capacity is limited, and that effective instructional design should take this into account. The theory focuses on how to manage the demands on working memory during learning to enhance comprehension and retention of information. According to CLT, there are three types of cognitive load that impact learning: *intrinsic load* (inherent complexity of the material being learned), *extraneous load* (cognitive load imposed by the way information is presented to learners), and *germane load* (cognitive load devoted to processing and integrating new information into existing knowledge structures).

A line of research has tested the cognitive load theory, highlighting the importance of reducing cognitive load and optimising the presentation of information to maximise learning outcomes (e.g., Mayer & Moreno, 2003; Paas & Sweller, 2003). Such empirical findings have several implications for instructional design: 1) breaking down complex topics into smaller, more manageable components to reduce intrinsic load, 2) removing irrelevant or redundant information to reduce extraneous load, and 3) combining visual and auditory elements (e.g., text, images, audio) to present information in a way that reduces cognitive load.

Vocabulary Ability and Story Comprehension

Vocabulary knowledge plays a critical role in reading comprehension. Past studies have investigated the relationship between vocabulary ability and story comprehension in children. For example, Elley and Mangubhai (1983) found that vocabulary size strongly predicted reading comprehension in New Zealand primary school children. Similarly, Ouellette (2006) found that children's vocabulary knowledge predicted their understanding of written narratives. The significance of morphological awareness—the capacity to identify and control the tiniest elements of meaning in words—has been the subject of recent research. Several studies have found that morphological awareness strongly predicts vocabulary and reading comprehension in children (e.g., Carlisle and Stone, 2005; Levesque et al., 2021; Nagy et al., 2006). However, the relationship between vocabulary ability and story comprehension is complex and may be influenced by other factors, such as background knowledge and cognitive processes (Cain, 2004; Perfetti & Stafura, 2014).

When children read or listen to a story, they attempt to decipher what the story tells. The story comprehension process forms a coherent mental representation in working memory (Mayer, 2009). This process involves at least three levels of representation. At the first “surface” level, children preserve the exact wording and rapidly fade from the memory. The second “textbase” level contains the text's meaning but not the exact wording. The third “situation” level represents the whole situation described and goes beyond the explicit text content. Children construct content-related connections based on their prior experiences and knowledge (Kinstch, 1998). Seger et al. (2021) examined these three levels of representation and found that they would be affected by adding illustrations to narrative text. Their findings showed that the reading times were significantly shorter when illustrations were presented than when not. However, the textbase was negatively affected when the illustration showed after its corresponding text. Besides, Wannagat et al. (2021) also provide evidence that audiovisual presentation is better than auditory presentation in promoting the coherence formation of representation and children's story comprehension.

Story comprehension involves cognitive skills that allow prereaders¹ to construct a coherent story representation. Empirical evidence showed that prereaders' vocabulary substantially affects story comprehension (Strasser & Río, 2014; Tōugu & Tulviste, 2017). In a storybook reading intervention for prereaders (preschoolers who could not read), Strasser, Larran, and Lissi (2013) discovered that those who heard the explicit meanings of five novel words per book understood the narrative better than those who did not. The result showed that improving vocabulary depth accelerates children's

story comprehension skills. In addition, Strasser & Rio (2014) also mentioned that vocabulary depth significantly mediates the effect of vocabulary breadth (size) on the story comprehension prereaders.

Story comprehension is an essential component of early literacy, as it requires the integration of various linguistic and cognitive skills (Kendeou et al., 2019). In a study of Malay-speaking children, Mohd Fadzil et al. (2021) found that children's story comprehension was influenced by their vocabulary knowledge and visual attention. Similarly, Zhang et al. (2020) reported that Chinese children with stronger vocabulary and visual attention skills demonstrated better story comprehension. However, more research is needed to understand the unique cultural and linguistic factors that influence story comprehension in Malaysian prereaders.

Eye-Tracking Research

Eye-tracking technology has been increasingly adopted to investigate visual attention, comprehension processes, and language development in young children (e.g., Justice et al., 2005; Hyönä, 2010; Rayner et al., 2006). Empirical studies have revealed various eye movement patterns that can serve as indicators of comprehension processes, such as selective revisiting behaviour (Rayner et al., 2006) and fixation duration on keywords (Evans & Saint-Aubin, 2005). Recent eye-tracking studies further confirm the principles of multimedia learning, showing that pictures and narration in storybooks increase children's visual attention, at once easing the comprehension process during reading (Justice et al., 2008; Kaefer et al., 2017; Verhallen & Bus, 2011). Visual attention plays a crucial role in prereaders' ability to process written texts and images (Fricke et al., 2016). In recent years, several studies have utilised eye-tracking technology to investigate the relationship between visual attention and reading ability in children (Blythe, 2014; Justice et al., 2019). For example, van den Boer et al. (2020) conducted two experiments to determine the extent to which young children fixate on the print of storybooks during shared book reading. They found that visual attention to print during shared reading was positively correlated with later reading outcomes in Dutch children.

While the aforementioned studies provide valuable insights, research on East Asian prereaders remains limited. Some studies explored young learners' literacy skills in the Malaysian context (e.g., Abdullah et al., 2019; Thang et al., 2021), but none of them used eye-tracking technology to track cognitive processes. While these studies did not use eye-tracking technology specifically, they do suggest that there may be a relationship between visual attention during reading and language skills in Malaysian children. It would be interesting to see if similar findings hold when eye-tracking technology is used to explore this relationship in more depth.

Using the eye-tracking methodology, researchers have shown that children integrate pictorial and sound information when attempting to understand a story. Wang (2020) further found that children started to pay attention to the protagonists of the picture book ages 2-3 years. In a read-aloud situation, children's reading comprehension levels highly correlate with their visual attention. (Takacs & Bus, 2018) provided evidence that children understand the story better on congruent pictures (pictures with a well-matched oral narration), whereas incongruent pictures cause a cognitive distraction in children. To comprehend the incongruent pictures, children tried to match the images they saw and the narration they heard but failed. Their eye movement results indirectly supported the dual-coding processing that both visual and verbal processing happen concurrently.

Liao et al. (2020) conducted an experimental study on a group of 4 to 5-year-old children to understand the children's visual attention during a shared storybook. At the pretest, the researchers found that children tend to look at the picture areas faster and longer than text areas when reading traditional read-aloud electronic storybooks. At the posttest, the conventional version was changed to two new designs to make print areas more attractive: one with highlighted text and one with text discussion. The eye-tracking results indicated that the two new storybook designs did attract children's attention more than the traditional read-aloud design. Additionally, Arslan-Ari & Ari (2021) demonstrated that children looked more at the images (82.31%) than at the texts (15.56%), and that 51.71% of fixations on text were synchronised with the visual stimuli and narration. In addition,

Cigerci and Gultekin (2017) examined the effects of digital stories on the listening comprehension of fourth-grade Turkish students and found that listening comprehension can be enhanced by viewing digital stories rather than simply listening to them.

Based on eye-tracking studies, the majority of prereaders rarely show interest in the printed text in picture storybooks. For instance, Evans & Saint-Aubin (2005) discovered that youngsters in picture storybooks spent less time fixating on the written text than the picture on each page, with longer text leading to longer fixations on the visuals rather than the written text. Justice et al. (2005) discovered that even while reading print-salient books, children only focused on the printed text for less than five seconds on average during the three minutes of reading time. It is crucial to remember that this convergent data comes from Western samples; therefore, it is still debatable if the findings can be generalised to a community in Southeast Asia, like Malaysia, where children are required to study many languages in school.

Early Childhood Education in Malaysia

With Malay as its official language and English and Chinese as its secondary languages, Malaysia is a multilingual and multicultural nation. English was originally the *de facto* official language throughout the British colonial expansion during the late 19th and early 20th centuries, before Malaysia attained independence in 1957. (Muniandy et al., 2010; Tan, 2013). In Malaysia, English is regarded as the second most important language after Malay, which is spoken by the majority of the country's population (Crystal, 1997, 2010; Tan, 2013; Muniandy et al., 2010). Malay and English were mandatory subjects in all elementary and secondary schools as a result of the educational policy established by the Ministry of Education in 1956 (also known as the Razak Report) (Gaudart, 1987; Tan, 2013). The country is committed to advancing a bilingual education system to foster bilingualism from an early age. Other commonly spoken languages in Malaysia include Malay dialects, Mandarin, Chinese heritage dialects, and Tamil. Malaysian children are generally exposed to many different languages from a young age, yet how this could affect their story comprehension and visual strategy is underexplored.

Children from different ethnic and language backgrounds may have varying experiences and challenges in early childhood education. Chiam and Lee (2015) investigated the effectiveness of a bilingual approach in early childhood education, which involved teaching in both English and Mandarin. The researchers found that children in the bilingual group showed more significant gains in language and literacy skills than those in the monolingual group. Moreover, children from different language backgrounds may have varying experiences in story comprehension. For example, children from Chinese-speaking backgrounds may have greater exposure to visual and spatial information, which may influence their comprehension of visual narratives (Leung and Ng, 2012). In contrast, children from Malay-speaking backgrounds may have greater exposure to oral storytelling traditions, which may influence their comprehension of oral narratives (Kumaran and Ramachandran, 2018).

Although all Malay and Chinese children tested in this study were born and lived in the same multi-ethnic country (i.e. Malaysia), their early literacy development might differ due to cultural differences in parental expectations (Keats et al., 1977; Masiran, 2022), formal education (Abidin et al., 2011), and home learning environment (Hill & Wagovich, 2020; Yeung & King, 2016). Malay and English are the two compulsory languages to be taught in preschool education for all Malaysian² children, especially in government preschools. However, private preschools offer more options to multi-ethnic groups using Malay, Chinese, Tamil, or other ethnic languages based on ethnicity and origin.

Since Malaysia is a multicultural country with a significant Muslim population, Islamic studies are often seen as necessary for Malay students' religious and cultural education. Malay preschools generally have 60 minutes of Islamic study weekly. In contrast, it is common for Chinese, Tamil, and other ethnic preschools to have 60 minutes of moral education weekly. Such moral education emphasises universal values, such as honesty, respect, and responsibility, which are relevant to all cultures and religions. Chinese preschools teach children three different languages – Chinese, English,

and Malay, compared to Malay preschool only offers two languages Malay and English (Ministry of Education Malaysia, 2017). In addition, Malay parents tend to emphasise religious-based values and train their children to read and memorise the Quran at a very young age (Bernama, June 17, 2016) whereas Chinese parents focus on mastering their children's literacy-related skills like print awareness and letter identification (Yeung & King, 2016).

The Current Study

The relationship between visual attention, vocabulary, and comprehension is intricate but essential in understanding early literacy development. With the eye-tracking technique, our study aims to elucidate these connections in the Malaysian context, especially among prereaders. This line of research can fill gaps in the literature of the field by providing insights into how these cognitive processes operate in different cultural and linguistic contexts so that educators and policymakers can develop more effective strategies to foster reading readiness. As aforementioned, Malaysia is a multilingual country, with Malay as the official language and English as a second language. Previous research has shown that bilingualism can influence cognitive processes, such as attention and working memory (e.g., Bialystok, 2007). Therefore, investigating these cognitive processes in Malaysian prereaders can provide valuable insights into how bilingualism affects cognitive development.

While previous studies have investigated this relationship in monolingual contexts, how it operates in a multilingual context remains unclear. Examining this relationship can provide insights into how children's vocabulary knowledge affects their comprehension of stories in a multilingual environment. The literature on story comprehension in children and its relation with visual attention lacks cohesiveness, and very few studies have been conducted in Southeast Asia. The intriguing questions remain unanswered: (1) How does the presentation of coherent narration along with picture and text content influence story comprehension in four- to five-year-old Malaysian prereaders? (2) To what extent does the presence of pictures and/or text aid understanding of narration and influence looking patterns in Malaysian prereaders?

The primary objective of this study was to investigate the effects of monomodal (auditory verbal information) versus multimodal (audiovisual) text presentation on the story comprehension performance of two groups of Malaysian children (aged 5–6 years) of two ethnic origins and from two educational backgrounds. While the children in this study had not received formal literacy instruction, we posited that the ambient linguistic environment, including exposure to different script systems, may yield systematic differences in visual processing strategies. Chinese children, who grow up in environments where they are exposed to logographic scripts that require holistic visual processing, might be hypothesized to allocate more attention to the pictorial element to gather contextual cues that aid in comprehending the story. Malay children, exposed to a linear and sequential script, might demonstrate a more balanced attention allocation between the text and the picture, or may even focus more on the textual component for comprehension. These distinct visual processing styles could potentially have a cascading effect on both story comprehension and vocabulary acquisition.

In the same vein, another aim was to examine whether the level of vocabulary skills modulates visual strategies for extracting information from the story stimuli. Eye tracking can measure where participants look while reading or listening to a story and how long they fixate on particular words or visual elements. This information can help researchers understand how visual attention influences prereaders' story comprehension and vocabulary acquisition, allowing for more accurate and detailed analysis than self-report measures. This is particularly important when studying young children who cannot accurately report their reading strategies or experiences.

The current is a follow-up of a study undertaken by the same research team (Sue et al., 2022). By using the same paradigm, our previous study showed that Chinese children in a kindergarten that uses English as the medium of instruction strongly preferred pictures over texts while listening to auditory narration; yet, they paid attention to the texts more when the narration was absent. The current study attempts to confirm the findings of Sue et al. (2022) and add scope and depth to the investigation by

comparing children from two different ethnic groups. This was done by adding information regarding the children's language and ethnic background and vocabulary skills. Based on Sue et al.'s (2022) findings and the previous literature on print-like static stories, we expected that, as compared to the text-and-narration-only condition, children would recall more details when encountering stories with illustrations that include congruent pictures that depict the same information in addition to the text and oral narration. More specifically, we anticipated that a congruent picture could lead to more optimal learning and effective looking strategies than an incongruent picture. Alternatively, the mental effort to manage the mismatching verbal and nonverbal information may interfere with listening to the story (Flack & Horst, 2018). If this prediction is accurate, children would recall fewer details from a story and revisit the objects and keywords more when an incongruent narration is presented simultaneously.

To examine if the different language backgrounds, educational environments, and cultural experiences affect prereaders' visual strategies for story comprehension, we compared the performance and eye gaze patterns between Malaysian-Malay and Malaysian-Chinese prereaders who are exposed to distinct educational traditions and cultural practices. It is crucial to emphasise that our intent is not to suggest inherent racial or biological differences in learning but to explore how varying cultural and educational backgrounds might influence the learning process. Also, our aim was not to portray Malaysian learners as an isolated or unique cognitive group. Instead, our investigation was rooted in the intricate linguistic landscape of Malaysia, where children often navigate bilingual or even trilingual environments from a tender age. Such rich linguistic exposure provides an opportunity to delve deeper into the interplay between language, cognition, and attention. Yet, our findings in this milieu are not strictly bound by national borders. The cognitive underpinnings we explored—how children allocate attention, process information, and construct meaning from narratives—are universally relevant. Indeed, as the world becomes increasingly interconnected and multilingual, many children across different countries will mirror the linguistic experiences of our Malaysian sample. Therefore, the insights garnered here, while contextualized within the Malaysian backdrop, hold broader implications for understanding cognitive processes in any multicultural or multilingual setting, affirming the universality of core cognitive features across diverse child populations.

Given that learning to read is a complex task that engages a range of cognitive operations, it is not surprising that it can pose a challenge for Malay and Chinese children who are learning English as an additional language while using an ethnic language at home. We anticipated that children's English proficiency indexed by their vocabulary skills is positively associated with story comprehension ability, and their looking strategies might differ depending on their language proficiency and ethnic background. By combining eye tracking with measures of vocabulary and comprehension, we aimed to gain a more comprehensive understanding of how visual attention contributes to these important language skills.

METHODS

Design

Similar to Sue et al. (2022), this study drew upon the research methodology developed by (Takacs & Bus, 2018), but it was modified to suit our research objectives. Each participant was assessed using a repeated-measures experimental design in four scenarios: 1) Text is presented with an accompanying picture and oral narration (congruent condition), 2) Text is presented with an accompanying picture, but the oral narration describes a scene that is not depicted in the text or the picture (incongruent condition), 3) Text is presented with an accompanying picture but no oral narration (control condition), and 4) Text is presented with matching narration but no picture (text condition). The details are summarised in Table 1.

Table 1. Characteristics of the four experimental conditions

Available Information	Congruent Condition	Incongruent Condition	Control Condition	Written Text Condition
Written text	√	√	√	√
Picture	√	√	√	
Oral narration	√	√		√

Participants

40 Malaysian preschool children (14 males and 26 females), aged between 5 to 6 years old ($M=5.65$; $SD=0.49$) were tested at two ethically-based kindergartens in Kuala Lumpur. 23 Chinese children were drawn from a kindergarten with only Chinese children. The parents of these children were mainly upper middle class with occupations such as lawyers, accountants, and engineers. The 17 Malay children were recruited from a kindergarten running with a Montessori programme (Montessori, 1912; Lillard, 2005) but with a strong emphasis on the Muslim religion. All the children in this kindergarten are Malays. The parents of these children range from middle to upper middle class with occupations such as construction workers, technicians, and teachers. The medium of instruction of both kindergartens was primarily English tested. The children ranged in age from 5 to 6 years old and had never received formal reading training. As a result, they were labelled as “prereaders.” They could speak and understand English but had received no formal English reading instruction. Before the research investigation began, parental approval was obtained.

Apparatus

To record eye movements, a Gazepoint GP3 HD eye tracker with a sampling rate of 150Hz and an accuracy of 0.5 was employed. The eye tracker works optimally at a distance of 65 cm from the eye. The eye-tracking technology is used in conjunction with two pieces of software: Gazepoint Analysis and Gazepoint Control. These tools were used to record fixation coordinates, establish AOIs, and generate metrics such as fixation durations and revisits. A regular chinrest located 60 cm from the monitor was used to limit participants’ head movements. Fixations on a preset key area of the stimulus were grouped to summarise the time and frequency an observer spent on specific locations when seeing the passage with or without visuals using the Areas of Interest (AOI) technique. The data on eye movements were then exported using Gazepoint Analysis software and analysed using IBM SPSS statistical software (version 28).

Materials

Four short English narratives with accompanying pictures, appropriate for children aged 5 to 6 years, were selected from four children’s storybooks. Before the study’s commencement, the selected stories were reviewed by a panel of experts in child literacy and cultural studies. These narratives were all similar in terms of difficulty and length (78-82 words) and were culturally appropriate for the children. They were used for the congruent condition and were modified slightly to assure that the five target objects were mentioned in-text.

As opposed to human-like figures, which infants tend to focus on more, the visual elements chosen were either objects or animals (Verhallen & Bus, 2011). To balance out salient and non-salient aspects, the images also had some components that were not mentioned in the narratives. Unlike in the case of the congruent condition, the four narratives for the incongruent condition used different sceneries and images despite having comparable primary characters. The research team created these narratives using the same standards –video length, degree of difficulty, and cultural appropriateness – compared to the congruent condition. A sample set of images and descriptions can be found in Appendix A.

The visual elements chosen were either objects or animals, but not human-like figures because children tend to pay more attention to humans, particularly to faces (Verhallen & Bus, 2011). The pictures also included some elements that were not mentioned in the narratives to have a balance of salient and non-salient details. Concerning the four narratives for the incongruent condition, although similar main characters were used, the scenery and images were different, unlike in the case of the congruent condition. These narratives were constructed by the research team following the same criteria as the narratives for the congruent condition, in terms of, length, difficulty level and cultural appropriateness. See Appendix A for a sample pair of pictures and narratives.

To evaluate children's English-language competence and measure their receptive vocabulary skills without requiring reading or writing, the Peabody Picture Vocabulary Test (Fourth Edition; Dunn & Dunn, 2007), which has high test-retest reliability and internal consistency, was adopted in the current study. This assessment tool also aids in evaluating language development among non-readers, even for individuals whose primary language is not English. This test has been widely used with monolingual English speakers and bilingual children in uninstructed settings (see Bialystok et al. 2010 for a meta-analysis). The use of the PPVT (or its British-English equivalent – the BPVS) is frequently used in L2 acquisition contexts as well, too (for instance, see Cohen 2016; Crevecoeur et al., 2014; Goriot et al., 2021). The test was administered verbally and took approximately 10-15 minutes. The children required no reading, and the examiner presented a series of pictures to each child. We selected the parallel test forms (Form A and Form B), each having 12 items for the difficulty levels of five and six-year-olds. There were four pictures on each page, and each was numbered. The examiner then stated a word describing one of the pictures and asked the child to point to or say the number of the picture that best represents the word. The total score was calculated based on the number of correct responses in the four-alternative-force-choice task.

Procedures

A 9-point calibration procedure was carried out at the start of the experimental activities to ensure eye movements were reliably registered with the eye tracker. Each child was then given a pair of headphones and instructed to gaze at the picture and/or text paragraph on the monitor while concurrently listening to the narration (except in the control condition where there was no narration). Each child was told to listen attentively to the narration coming via the headphones since he/she would need to remember it later. Each trial was displayed for 30 seconds. To encourage participants to rely on their long-term memory when recalling the story, the child performed a short distracter task involving solving an object problem or building a Lego toy once the narration was finished. Subsequently, a facilitator would ask the youngster to recollect the tale he or she heard after the game. The narration of the narrative was audio-recorded. The game was added so the children would remember the story based on what they had learned rather than just repeating the words retained in their short-term memory. The procedure took about 30 minutes and was carried out four times based on the pseudo-counterbalanced order of conditions shown in Table 1. After the storytelling tasks, participants completed the Peabody Picture Vocabulary Test (Fourth Edition), which measured their receptive vocabulary.

Data Analysis

The audio recordings of storytelling were transcribed verbatim. Then, a standardised rating rubric (see Appendix B) was used to score how well the children recalled the elements of the stories they heard based on a 5-point scale whereby 0 represents no recall of the story, and 4 represents a complete summary of all the events of the story. Three independent raters did the ratings to minimise experimenter bias. In any case of discrepancy, a discussion was undertaken to arrive at a consensus. In the preliminary analyses, the storytelling performance and English-language proficiency were first analysed separately for the two groups of children – Chinese versus Malays – to check if their dissimilar language backgrounds are associated with story comprehension and vocabulary skills. Two mixed factorial ANOVAs, including the storytelling performance (good, average, or weak) and vocabulary

level (low, medium, high) as between-subject variables, and the AOIs (image, paragraph, objects, and keywords) and display conditions as the within-subject variables, were also performed on the eye movement indexes: the average proportional fixation duration and the average number of revisits. The average proportional fixation duration represents how long the average fixation lasted on each predefined AOI. In the context of story comprehension, it might indicate that participants are actively trying to process and integrate the information in those AOIs into their understanding of the story. The number of revisits shows how often a participant looked back at a certain area of interest (AOI). This allows the researcher to see which parts of the display held the participant's attention (for better or worse) and which parts they quickly moved on from. Revisits to resolve incoherent information can arguably reflect switching from automatic to deliberate processes during story comprehension.

RESULTS

Story Comprehension Scores

A 2 (school type: Chinese and Malay) \times 4 (display conditions) repeated-measure ANOVA revealed a marginally significant main effect of display conditions, $F(2.41, 82.00) = 2.38, p = 0.07, \eta_p^2 = .52$, where children generally performed better in the congruent condition than the incongruent ($p = .02$), control ($p = .01$), and text ($p = .08$) conditions. Neither the main effect of school type (Chinese vs. Malay), $F(1, 34) = 1.42, p = .24, \eta_p^2 = .04$, nor the interaction, $F(3, 102) = 1.90, p = .14, \eta_p^2 = .05$, was significant. The following table shows the descriptive values for both groups of children.

Vocabulary Scores

A 2 (school type: Chinese and Malay) \times 2 (test difficulty: 5- and 6-year-old) repeated-measure ANOVA performed on the vocabulary scores revealed a significant main effect of test difficulty, $F(1, 35) = 12.52, p = .001, \eta_p^2 = .26$. Regardless of the school type, children made more errors in the 6-year-old version of the vocabulary test (Chinese: $M = 3.10, SE = 0.58$; Malay: $M = 4.63, SE = 0.73$) than in the 5-year-old version of the vocabulary test (Chinese: $M = 1.76, SE = 0.49$; Malay: $M = 2.69, SE = 0.73$) of the receptive language proficiency test, indicating the increased task difficulty introduced by more unique objects or less culturally familiar items. However, the main effect of school type, $F(1, 35) = 2.65, p = .11, \eta_p^2 = .07$, and the interaction, $F(1, 35) = 0.43, p = .52, \eta_p^2 = .01$, failed to reach statistical significance.

Inter-Task Correlations

Pearson correlation tests (two-tailed, $\alpha = 0.05$) showed a significant positive relationship between story comprehension and vocabulary skills in the children studying at the Malay kindergarten, $r(17) = .60, p = .01$, but not in those from the Chinese kindergarten, $r(19) = .23, p = .37$.

Table 2. Mean scores and standard errors of the mean of different story display conditions among children from Chinese and Malay kindergartens

	Chinese		Malay		Overall	
	M	SE	M	SE	M	SE
Congruent	1.32	0.28	1.94	0.33	1.61	0.22
Incongruent	1.16	0.18	1.24	0.18	1.19	0.13
Control	1.16	0.22	1.12	0.24	1.14	0.16
Text	0.89	0.3	1.65	0.41	1.25	0.25

Average Proportional Fixation Duration

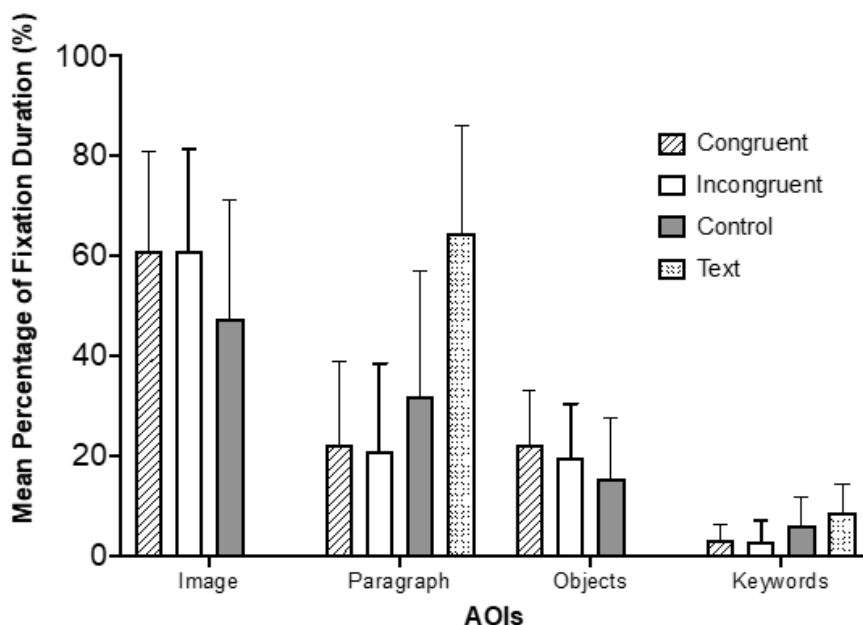
A repeated-measures ANOVA revealed a significant main effect of condition, $F(3, 57) = 14.21$, $p < .001$, $\eta_p^2 = 0.43$, and a significant main effect of AOI, $F(3, 57) = 45.95$, $p < .001$, $\eta_p^2 = 0.71$. Overall, children spent less time looking at the stimuli in the text-only condition than in the congruent, incongruent and control conditions (all $ps < .001$). For the AOIs, children looked longer at the image than the text paragraph ($p = 0.03$); more specifically, they fixated longer on the target objects than the keywords ($p < .004$), indicating that the children were paying more attention to the pictorial information. The interaction between condition and AOI also reached significance, $F(9, 171) = 51.47$, $p < .001$, $\eta_p^2 = 0.73$. Post hoc comparisons revealed that both incongruent and congruent images received longer fixations than the control image (both $p < .001$). As expected, the paragraph in the text-only condition received longer looking times than all other conditions (all $p < .001$). In the control condition, children looked longer at the text paragraph than in both incongruent ($p = 0.018$) and congruent ($p = 0.033$) conditions. However, for objects and keywords, no between-condition difference was found (all $p = 1$).

There was a significant interaction between display condition and storytelling performance, $F(6, 57) = 3.25$, $p = .008$, $\eta_p^2 = .26$, accompanied by a significant three-way interaction between the display condition, story comprehension, and school type, $F(6, 84) = 3.14$, $p = .008$, $\eta_p^2 = .18$. Pairwise comparisons revealed that high-performing children from the Malay school looked longer at the display than those from the Chinese school, particularly in the congruent ($p = .03$), incongruent ($p = .04$), and text condition ($p = .02$), but not in the control condition ($p = .57$). The main effect of vocabulary level and all the interactions related to it were non-significant (all $ps < .05$).

Average Number of Revisits

A repeated-measures ANOVA performed on the average number of revisits revealed a significant main effect of AOI, $F(3, 108) = 51.73$, $p < .001$, and a significant main effect of condition, $F(3, 108) = 4.00$, $p = 0.01$. The image ($M = 42.43$, $SE = 14.50$) received the greatest number of revisits, followed by paragraph ($M = 34.91$, $SE = 10.21$), objects ($M = 14.31$, $SE = 4.98$), and keywords ($M = 5.05$, $SE = 1.34$).

Figure 2. The mean proportional fixation duration (in percentages) on each AOI for different presentation conditions



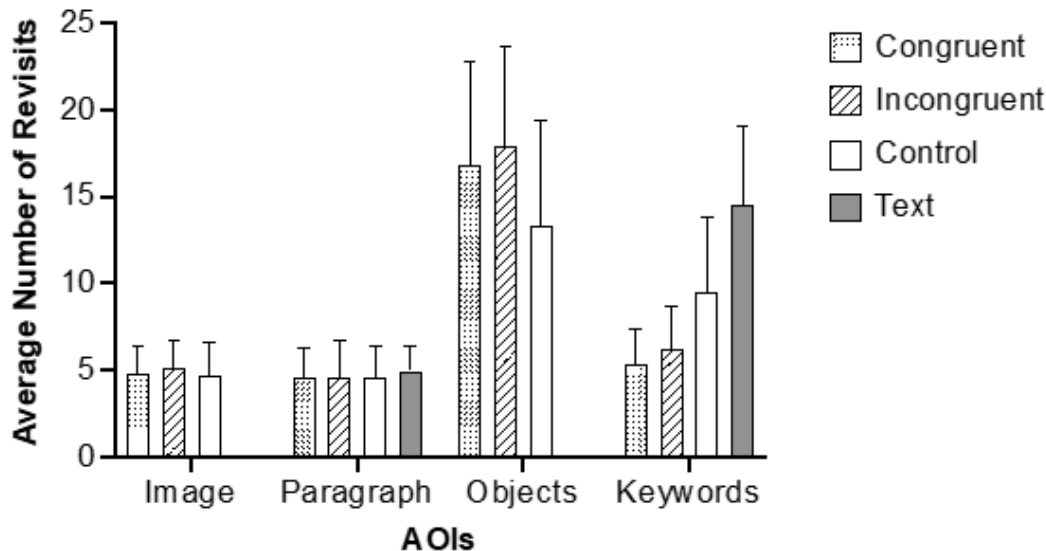
(all $ps < .05$). When comparing between conditions, children generally made more revisits in the text-only condition ($M=9.79$, $SE=4.79$) than in the congruent ($M=7.94$, $SE=2.98$), incongruent ($M=8.47$, $SE=3.15$), and control ($M=8.06$, $SE=3.15$) conditions. However, the interaction was marginally significant, $F(9,324)=2.11$, $p=0.06$. Simple main effect analysis revealed no difference between conditions for AOIs (all $p > .05$), except for the keywords. Children revisited the keywords in the text-only conditions more than all other conditions (all $ps \leq 0.008$).

DISCUSSION

The development of early literacy skills is a critical aspect of a child's cognitive development. Among these skills, visual attention, story comprehension, and vocabulary are interrelated components that contribute to a child's overall reading ability. Eye-tracking technology has been instrumental in investigating these relationships in young children, as it provides objective measures of visual attention during the reading process. The current eye-tracking study investigated the relationship between visual attention, story comprehension, and vocabulary skills in Malaysian prereaders. To test if children make an association between narration and visual images during storybook reading, we also recorded their eye movements – average proportional fixation duration and average number of revisits – which could reflect the processing of relevant/irrelevant visual depictions. The average fixation period was the longest for the image when comparing AOIs across all the pictorial conditions, suggesting that using pictures in storybooks is beneficial for grabbing visual attention. The eye-tracking data also showed that prereaders were more engaged when an image was present, regardless of whether the image matched the narration or not. In line with Montag et al. (2015), which found that the presence of pictures helps children focus better on the story, our result indicates that Malaysian prereaders had a greater interest in and understanding of the story threads when presented through pictures.

The eye-tracking data results showed that pictures attract the visual attention of Malay and Chinese prereaders for story comprehension, irrespective of their ethnic origins and educational background. The study findings are consistent with previous research (e.g., Evans et al., 2009;

Figure 3. The average number of revisits on each AOI for the four different presentation conditions



Montag et al., 2015; Van Der Schoot et al., 2008), indicating that Malaysian prereaders were more interested in and had a better understanding of storylines presented with pictures rather than only text. The prereaders paid less attention to the story without a picture, as evidenced by their shorter fixation durations on the text stimuli. The study also found that prereaders were not distracted by a non-matching picture in the incongruent condition, and the presence of narration and a matching picture helped them focus better on the story. The results further suggest that the prereaders' overall attentiveness when listening to stories was higher with the presence of both text and pictures compared to text alone. These findings align with the Cognitive Theory of Multimedia Learning (Mayer, 2005), which proposes that deeper learning can occur when information is presented in both text and graphics, rather than just text.

In the current study, we found no clear evidence that different language backgrounds between Malay and Chinese prereaders directly affect their story comprehension ability, receptive vocabulary skills, and looking patterns. Despite a mismatched narration, prereaders continued to search for objects in a picture. Additionally, there was no notable disparity in the prereaders' visual interest in keywords between the congruent and incongruent conditions. Even in the text-only condition, where the prereaders spent more time fixating on the text, they did not have significantly longer fixation durations on the keywords. However, children revisited the keywords in the text-only conditions more than all other conditions. Consistent with the notion that selective revisiting behaviour is a reliable eye-movement indicator of deliberately resolving incoherent information, the prereaders might not have reached the level at which they could automatically associate the keywords with the narration or the objects in the pictures. Alternatively, it could be the case that keywords are not an essential consideration among prereaders who are still developing print awareness. Follow-up studies should be conducted on children of a higher age range to rule out one of these seemingly plausible explanations.

Concerning children's language proficiency, our comparisons between low, medium, and high-performing children in the story comprehension task highlighted the importance of considering storytelling ability as well as their language background when investigating the visual strategies of children. It is important to note that the current study was not designed to systematically capture qualitative differences in concretising audiovisual narration between Malay and Chinese children. However, our results showed that high-performing Malay children looked longer at the displayed stimuli than high-performing Chinese children, particularly in the three conditions – congruent, incongruent, and text – where a narration was presented along with the visual stimulus. Such a pattern suggests a stronger reliance on simultaneous visual and narrative information processing in short-term memory. This finding is consistent with research on the use of multimedia in education, which has shown that the use of visual and audio aids can enhance learning and memory (Mayer, 2005). Additionally, Bernama (2016) reported that Malay children are trained to read and memorise the Quran with visual and audio aids at a young age. Hence, this suggests that cultural variations in language development and mental representation may drive the interaction between ethnic groups and looking strategies. Further research would be needed to explore this possibility.

Past studies have shown that children from different language backgrounds may have varying experiences and challenges in early education (Chiam & Lee, 2015). In Malaysia, Mandarin is often used as a medium of instruction in Chinese kindergartens, while Malay is used in Malay kindergartens. Such a difference may have influenced the language exposure and vocabulary development of the children from the two kindergartens differently. Moreover, research has also shown that language exposure and use can influence the development of cognitive processes such as attention and working memory (Bialystok, 2007). Therefore, the cognitive processes involved in story comprehension might differ between the two groups of children based on their language backgrounds and exposure. In fact, cultural and linguistic differences can influence how individuals process visual information (Masuda & Nisbett, 2001), and cultural variations in language learning

environments can affect cognitive development (Kuhl, 2010). Future research is required to test this hypothesis, along with additional consideration of vocabulary proficiency and verbal fluency (Perfetti et al., 2005). Overall, these findings imply the importance of considering children's diverse language and cultural backgrounds in early education and tailoring instruction and assessment to meet their needs.

Regardless of differences in ethnic backgrounds and language proficiency, providing pictorial illustrations to Malaysian children while listening to narrations seems crucial. Moreover, these results challenge the prevailing trends of audiobooks and story podcasts, such that sounds alone stimulate young children more than pictures. Our findings have practical implications for creating digital picture storybooks (e.g., implementing additional stimulus manipulations to promote the integration of the narrative and nonverbal information when listening to storybooks) in different cultural settings. During the COVID-19 pandemic, our prereaders were required to listen to online lessons and programmes regularly; therefore, incorporating these improved digital storybooks into their online classes could make their learning more fun and enhance their listening and story comprehension skills, enabling them to learn new vocabulary.

In animated storybooks, Takacs and Bus (2016) showed that animated stories more easily attract 4-6-year-old children's attention than static pictures. Animated pictures that are well-matched to the text of the story most probably direct children's attention to specific details highlighted by the story text. The motion in images changes the way children process pictorial information. Due to motion, children moved less between the different visual elements and fixated more on the key images, resulting in a more in-depth exploration and better integration of verbal and visual information. This explains why children tend to look longer at animated books than static ones, consequently elevating story comprehension. Given that animations attract visual attention and change how children process illustrations, further research could examine how e-storybooks could utilise dynamic stimuli to reinforce the beneficial effects of storybook reading.

The current findings have a few theoretical implications. First, it strengthens the proposition of the cross-channel connections between auditory and pictorial representations in enhancing story comprehension and vocabulary learning. Second, the finding that both Malay and Chinese prereaders displayed similar patterns of visual attention, especially when it comes to the absence of interest in printed text alongside pictures, suggests that the early stages of literacy development may have universal elements that are not heavily influenced by ethnic origins. Third, the shift in attention from images to text in the absence of narration suggests that prereaders attempt to decipher the meaning of unfamiliar symbols (text) when auditory cues are missing. This resonates with cognitive theories proposing that learners allocate attentional resources based on task demands. Fourth, the differential patterns of attention to target objects and keywords based on storytelling performance underscore the role of pre-existing knowledge and comprehension abilities in guiding visual attention.

CONCLUSION

In summary, our research indicated that the inclusion of any picture significantly enhances the visual attention of Malaysian preschoolers to brief stories and storytelling performances. The eye-tracking data also revealed that high-performing prereaders may use distinct viewing techniques to enhance comprehension. Having a picture present boosts attention compared to having no visual stimulus at all. Even in cases where there was no oral narration or when the picture did not correspond with the narration, prereaders' eye movements indicate a significant effort to extract information from the stimuli with the presence of an image. This implies that images are a powerful tool to guide children's visual attention to particular details. Mayer's Cognitive Theory of Multimedia Learning suggests that effective early literacy instruction should consider integrating visual and verbal information to support the development of mental representations and promote comprehension. Further research is

required to investigate how different teaching methods and language backgrounds affect children's story comprehension ability in Malaysia. Examining these factors can provide insights into how early childhood education can be tailored to meet the diverse needs of children from different cultural and linguistic backgrounds.

COMPETING INTERESTS

The authors of this publication declare they have no competing interests.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the author(s) of the article.

REFERENCES

- Abidin, M. J. Z., Rezaee, A. A., Abdullah, H. N., & Balbir Singh, K. K. (2011). Learning styles and overall academic achievement. *International Journal of Humanities and Social Science*, 1(10).
- Ann Evans, M., & Saint-Aubin, J. (2005). What Children Are Looking at During Shared Storybook Reading: Evidence From Eye Movement Monitoring. *Psychological Science*, 16(11), 913–920. doi:10.1111/j.1467-9280.2005.01636.x PMID:16262779
- Arslan-Ari, I., & Ari, F. (2021). The effect of visual cues in e-books on pre-K children's visual attention, word recognition, and comprehension: An eye tracking study. *Journal of Research on Technology in Education*, 1–15. doi:10.1080/15391523.2021.1938763
- Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual Differences in the Inference of Word Meanings From Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology*, 96(4), 671–681. doi:10.1037/0022-0663.96.4.671
- Cigerci, F. M., & Gultekin, M. (2017). *Use of digital stories to develop listening comprehension skills*. Academic Press.
- Cohen, C. (2016). Relating input factors and dual language proficiency in French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, 19(3), 296–313. doi:10.1080/13670050.2014.982506
- Crevecoeur, Y. C., Coyne, M. D., & McCoach, D. B. (2014). English language learners and English-only learners' response to direct vocabulary instruction. *Reading & Writing Quarterly*, 30(1), 51–78. doi:10.1080/10573569.2013.758943
- Crystal, D. (1997). *English as a Global Language*. Cambridge University Press.
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). APA PsycTests [Database record]. doi:10.1037/t15144-000
- Evans, M. A., Saint-Aubin, J., & Landry, N. (2009). Letter Names and Alphabet Book Reading by Senior Kindergarteners: An Eye Movement Study. *Child Development*, 80(6), 1824–1841. doi:10.1111/j.1467-8624.2009.01370.x PMID:19930354
- Flack, Z. M., & Horst, J. S. (2018). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development*, 27(1), e2047. doi:10.1002/icd.2047
- Gaudart, H. (1987). A typology of bilingual education in Malaysia. *Journal of Multilingual and Multicultural Development*, 8(6), 529–552. doi:10.1080/01434632.1987.9994310
- Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2021). Using the peabody picture vocabulary test in L2 children and adolescents: Effects of L1. *International Journal of Bilingual Education and Bilingualism*, 24(4), 546–568. doi:10.1080/13670050.2018.1494131
- Hill, M. S., & Wagovich, S. A. (2020). Word learning from context in school-age children: Relations with language ability and executive function. *Journal of Child Language*, 47(5), 1006–1029. doi:10.1017/S0305000919000989 PMID:32252839
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. doi:10.1037/0033-295X.87.4.329 PMID:7413885
- Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology*, 44(3), 855–866. doi:10.1037/0012-1649.44.3.855 PMID:18473649
- Justice, L. M., Skibbe, L., Canning, A., & Lankford, C. (2005). Pre-schoolers, print and storybooks: An observational study using eye movement analysis. *Journal of Research in Reading*, 28(3), 229–243. doi:10.1111/j.1467-9817.2005.00267.x

- Kaefer, T., Pinkham, A. M., & Neuman, S. B. (2017). Seeing and knowing: Attention to illustrations during storybook reading and narrative comprehension in 2-year-olds. *Infant and Child Development*, 26(5), e2018. doi:10.1002/icd.2018
- Keats, D. M., Keats, J. A., & Abdul Rahman, W. R. (1977). Performance of Malay and Chinese children on vocabulary and operational tasks in two languages. *Australian Psychologist*, 12(1), 63–68. doi:10.1080/00050067708255861
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2021). How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development. *Journal of Research in Reading*, 44(1), 10–26. doi:10.1111/1467-9817.12313
- Liao, C.-N., Chang, K.-E., Huang, Y.-C., & Sung, Y.-T. (2020). Electronic storybook design, kindergartners' visual attention, and print awareness: An eye-tracking investigation. *Computers & Education*, 144, 103703. doi:10.1016/j.compedu.2019.103703
- Lillard, A. S. (2005). *Montessori: The science behind the genius*. Oxford University Press.
- Masiran, R. (2022). A Review of Parenting in A Multicultural Country: The Malaysian Experience. In S. Ali Samadi (Ed.), *Parenting—Challenges of Child Rearing in a Changing Society*. IntechOpen. doi:10.5772/intechopen.101173
- Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (1st ed., pp. 31–48). Cambridge University Press. doi:10.1017/CBO9780511816819.004
- Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *The American Psychologist*, 63(8), 760–769. doi:10.1037/0003-066X.63.8.760 PMID:19014238
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52. doi:10.1207/S15326985EP3801_6
- Ministry of Education Malaysia. (2017). *Dasar Pendidikan Kebangsaan*. https://www.moe.gov.my/images/KPM/UKK/2019/03_Mac/BUKU_DPK_JILID_KEEMPAT.pdf
- Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The Words Children Hear: Picture Books and the Statistics for Language Learning. *Psychological Science*, 26(9), 1489–1496. doi:10.1177/0956797615594361 PMID:26243292
- Montessori, M. (1912). *The Montessori Method: Scientific Pedagogy as Applied to Child Education in the Children's Houses*. Frederick A. Stokes.
- Muniandy, B. (2010). Academic Use of Internet among Undergraduate Students: A Preliminary Case Study in a Malaysian University. *International Journal of Cyber Society and Education*, 3, 171–178.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4. doi:10.1207/S15326985EP3801_1
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. L. Erlbaum Associates.
- Perfetti, C., & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, 18(1), 22–37. doi:10.1080/10888438.2013.827687
- Rudolph, M. (2017). *Cognitive Theory of Multimedia Learning*. Academic Press.
- Seger, B. T., Wannagat, W., & Nieding, G. (2021). Children's surface, textbase, and situation model representations of written and illustrated written narrative text. *Reading and Writing*, 34(6), 1415–1440. doi:10.1007/s11145-020-10118-1
- Sorden, S. D. (2013). The Cognitive Theory of Multimedia Learning. In *The Handbook of educational theories* (pp. 155–168). Information Age Publishing, Inc.
- Strasser, K., & del Río, F. (2014). The Role of Comprehension Monitoring, Theory of Mind, and Vocabulary Depth in Predicting Story Comprehension and Recall of Kindergarten Children. *Reading Research Quarterly*, 49(2), 169–187. doi:10.1002/rtrq.68

Strasser, K., Larraín, A., & Lissi, M. R. (2013). Effects of storybook reading style on comprehension: The role of word elaboration and coherence questions. *Early Education and Development*, 24(5), 616–639. doi:10.1080/10409289.2012.715570

Sue, C. H., Thang, S. M., Wong, H. K., Tan, P. S., Loo, F. L., & Ahju, R. (2022). To what extent do pictures support Malaysian prereaders' visual attention and retelling of stories: An eye-tracking experiment. *International Journal of Computer-Assisted Language Learning and Teaching*, 12(4), 1–17. doi:10.4018/IJCALLT.310080

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295–312. doi:10.1016/0959-4752(94)90003-5

Takacs, Z. K., & Bus, A. G. (2016). Benefits of Motion in Animated Storybooks for Children's Visual Attention and Story Comprehension. An Eye-Tracking Study. *Frontiers in Psychology*, 7. Advance online publication. doi:10.3389/fpsyg.2016.01591 PMID:27790183

Takacs, Z. K., & Bus, A. G. (2018). How pictures in picture storybooks support young children's story comprehension: An eye-tracking experiment. *Journal of Experimental Child Psychology*, 174, 1–12. doi:10.1016/j.jecp.2018.04.013 PMID:29857060

Tan, S. I. (2013). *Malaysian English: Language Contact and Change*. Peter Lang. doi:10.3726/978-3-653-03516-2

Thang, S. M., Jaafar, N. M., Ho, H. F., & Abdul Aziz, N. B. (2021). Using Eye Tracking to Investigate Strategies Used by ESL Learners in Reading a Scientific Text with Diagram. In L. Miller & J. G. Wu (Eds.), *Language Learning with Technology*. Springer. doi:10.1007/978-981-16-2697-5_15

Tōugu, P., & Tulviste, T. (2017). Links between children's language comprehension and expressive vocabulary and their personality traits: A longitudinal investigation. *Learning and Individual Differences*, 57, 123–128. doi:10.1016/j.lindif.2017.03.015

Van Der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. D. M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading*, 31(2), 203–223. doi:10.1111/j.1467-9817.2007.00354.x

Verhallen, M. J. A. J., & Bus, A. G. (2011). Young second language learners' visual attention to illustrations in storybooks. *Journal of Early Childhood Literacy*, 11(4), 480–500. doi:10.1177/1468798411416785

Wagovich, S. A., Anderson, J. D., & Hill, M. S. (2020). Visual exogenous and endogenous attention and visual memory in preschool children who stutter. *Journal of Fluency Disorders*, 66, 105792. doi:10.1016/j.jfludis.2020.105792 PMID:33032169

Wang, J. (2020). Children's visual attention characteristics reveal the level of information picture book reading comprehension. *Journal of Chinese Writing Systems*, 4(1), 27–44. doi:10.1177/2513850219888029

Wannagat, W., Waizenegger, G., & Nieding, G. (2021). Coherence formation during narrative text processing: A comparison between auditory and audiovisual text presentation in 9- to 12-year-old children. *Cognitive Processing*, 22(2), 299–310. doi:10.1007/s10339-020-01008-z PMID:33404902

Yeung, S. S., & King, R. B. (2016). Home Literacy Environment and English Language and Literacy Skills Among Chinese Young Children Who Learn English as a Second Language. *Reading Psychology*, 37(1), 92–120. doi:10.1080/02702711.2015.1009591

ENDNOTES

- ¹ “Prereaders” is a term used to refer to children who are in the developmental stage before they have acquired full reading proficiency. These are typically young children who are in the process of learning the foundational skills and concepts necessary for reading. Prereaders are at the early stages of literacy development and are often characterized by their exposure to language, engagement with books and stories, and the gradual acquisition of pre-reading skills such as phonological awareness, letter recognition, and understanding basic story structures. The prereading stage is crucial as it lays the groundwork for later reading success, enabling children to transition from recognizing individual words to comprehending and engaging with more complex texts.

- ² The term “Malaysian” pertains to the larger sociocultural context of the country, encompassing its diverse ethnic, linguistic, and cultural makeup. On the other hand, “Malay” refers to one of the major ethnic groups in Malaysia, with its own distinct language and cultural identity. Throughout this study, we alternately refer to both the “Malaysian” context, representing the broader societal landscape, and the “Malay” population, reflecting the specific linguistic and ethnic group under examination. This distinction aims to ensure clarity and precision in discussing the intricacies of our research within the Malaysian context, while simultaneously acknowledging the diverse linguistic and cultural dimensions that constitute Malaysia as a nation.

Siew Ming Thang is a professor at ELM Graduate School, HELP University. Her areas of interest are Technologies in Education, Learner Autonomy, Motivation and Learning and Eyetracking Research.

Loo Fung Lan is a lecturer at Open University Malaysia. She is also the programme director for the Bachelor of Early Childhood Education programme. She was the founder and principal of a private kindergarten and a learning centre for more than 12 years. Through the years, she was passionate about educating young children. She obtained a master's degree in Early Childhood Education from the University of National Malaysia and a PhD from the University of Malaya. She specializes in early childhood education and her areas of interest are early literacy, child psychology, pedagogy, and curriculum development.