Is There a Relationship between Prereaders' Visual Attention and Their Storytelling Performance? Evidence from Eye-Tracking and Qualitative Data

Siew Ming Thang (tsm.2001uk@gmail.com) *Corresponding Author
HELP University, Malaysia

Muthukrishnan Priyadarshini (priyadarshini.m@help.edu.my
HELP University, Malaysia

Jennifer Poh Sim Tan (jennifer.tan@help.edu.my)
HELP University, Malaysia

Hoo Keat Wong (HooKeat.Wong@nottingham.edu.my)
(University of Nottingham, Malaysia)

Arshad Nurul Iman (NurulIman.Arshad@taylors.edu.my)
(Taylor's College, Malaysia)

Chee Hao Sue (suecheehao@gmail.com)
(HELP University, Malaysia)

Abstract

In the Malaysian context, numerous studies explored young learners' literacy skills but none uses the eye-tracking device to track the cognitive processes of prereaders when they are reading picture storybooks. For this project 22 prereaders (4-5 years old) listened to brief stories in four conditions: (a) only static text with narration (text condition), (b) oral narration, and a picture that was congruent with the narration (congruent condition), (c) oral narration and an incongruent picture (incongruent condition), and (d) only picture with text but no oral narration (control condition). Inferential statistics were used to analyse the relationship between eye fixation, gender, and story-telling performance. The quantitative analysis revealed that the prereaders' story-telling performance did not depend on treatment conditions, gender, or fixation patterns on the predefined areas of interest (AOI). To further understand this phenomenon a qualitative analysis was undertaken to explore the patterns of the prereaders' storytelling performance (STP). The qualitative analysis suggested affective factors such as the prereaders' state of mind and their level of attentiveness could have affected their STP. Hence, they may appear to be focusing on the screen but their attention could have been somewhere else or they could have been confused. In addition, their lack of proficiency in a language that was not their first/native language and their general inability to communicate well verbally could have been contributory factors.

Introduction

The use of the eye-tracker to explore young children's cognitive processes has increased in popularity in recent years (e.g. Evans & Saint-Aubin, 2005; Verhallen & Bus, 2011; Takacs & Bus, 2016, 2018). Cognitive processes refer to the tasks the brain does continuously in processing all the information received from the environment. One of the key cognitive processes is attention which makes it possible for us to position ourselves towards certain stimuli and respond to them. Visual attention can be investigated through eye-tracking research through eye movements and fixation patterns.

One particularly interesting study that probed into young children's cognitive processes was conducted by Takacs and Bus (2018) using the eye-tracker as a research method. It investigated prereaders reading picture storybooks while listening to a narration. Prereaders are defined as young children who have not received formal reading instructions. The reason why they conducted a study on the utilisation of both skills was that eye-tracking studies have generally revealed that pre-readers rarely show interest in the printed text available in most picture storybooks in comparison to the picture on the same page. For example, Evans and Saint Aubin (2005) found that children spent on average 175-669 milliseconds fixating on the written text per page in a picture storybook versus 7,235-18,605 milliseconds fixating on the pictures. In the case of longer text, they would spend even more time looking at the pictures but not more time on the written text. This study is supported by more recent research which revealed that preliterate children seldom glance at the text during storybook reading (Gong & Levy, 2009; Justice et al., 2005, 2008; Liu, 2011; Liao et al., 2013, 2020; Lin et al., 2018). Hence, Takacs and Bus (2018) were interested to find out whether a combination of listening and reading activities would make a difference in the amount of time the children spent on the picture and text.

Takacs and Bus's (2018) study was conducted on 41 children (from 4 to 6 years old) listening to a text in Dutch (which is their first language) with a storybook format under four different conditions i.e. (a) only static text with narration (text condition), (b) oral narration and a picture that was congruent with the narration (congruent condition), (c) oral narration and an incongruent picture (incongruent condition), and (d) only picture with text but no oral narration (control condition). The children's eye movements were recorded with a Tobii eye-tracker with 3.2.1 software. After that, the children were given a game to play. This was to assure that the children would not just repeat what was stored in their short-term memory. Following that the children were asked to tell the story they heard. For condition 3, where there was no narration, they were asked to describe the picture they saw. An analysis of variance (ANOVA) with repeated measures for story recall per condition for the four conditions was carried out. The results revealed that the children's narration was more similar to the original narration when they were allowed to hear the narration (conditions a, b, and c), as compared to the picture control condition (d). This suggested that the picture alone was not sufficient for the children to make up the story. The results further showed that the children told the story better in the congruent condition (b) than in the text condition and the incongruent condition. There was no difference between incongruent and text conditions. Thus, it appears that children's retelling was substantially increased in the situation where congruent narration and pictures were present which suggested the advantage of interactive reading and picture storybooks. Although retelling performance decreased when the incongruent

narration was used, it did not have a detrimental effect on story comprehension.

The current study took their study a step further by undertaking a qualitative inquiry to find out whether there is a relationship between the patterns and quality of their storytelling and the conditions they were exposed to. These qualitative data were further triangulated with the fixation patterns derived from the eye-tracking analysis in an attempt to understand better the relationship between eye movements, listening skills, and story-telling performance (STP).

Although numerous studies have been undertaken in the Malaysian context to explore young learners' literacy skills, none of them have used the eye-tracking device to track their cognitive processes while they are reading and listening at the same time. This type of study is crucial as it enables us to understand the relationship between these two skills better which will lead to the development of more suitable reading materials for young children. In addition to that, it informs educators on why it is important to improve storytelling performance and suggests ways to go about it.

Literature Review

Relevant Theories

Many theories focus on text-picture comprehension (TPC) but three theories specifically target formats of mental representations involved in TPC. They are:

(1) Dual Coding Theory (Paivio, 1986),
(2) The Cognitive Theory of Multimedia Learning (Mayer, 2005) and
(3) The Integrative Model of Text Picture Comprehension (TPC) (Schnotz & Bannert, 2003)

The three theories share the assumption of separate channels for text and picture processing but differ on other issues. Paivio's (1971, 1990) DCT emphasizes two processing channels for verbal and non-verbal information which are functionally independent, yet interconnected. The verbal system processes verbal information, such as spoken or written words, regardless of the modality of origin while the nonverbal system processes nonverbal information, such as pictures, gestures, and music, again, regardless of origin. These two systems thereby enable simultaneous processing of visual and narrative information in short-term memory. Paivio's dual-coding theory (DCT) was used to hypothesise that pictures in storybooks facilitate understanding of the narration, more so when they were closely matched to the narration (Mayer, 2005).

The current study further draws upon applied linguistics research that propose visuals influence the listening comprehension process. For example, studies in L1 acquisition have shown, for instance, that children rely heavily on non-verbal cues when developing their L1 speech perception (Ochs & Schieffelin, 2009). Similarly, research on L2 listening comprehension has evinced that visuals can facilitate the development of L2 listening skills.

Visually enhanced L2 listening comprehension can be better understood through the following two theoretical perspectives and their underlying assumptions. Mayer's (2005) cognitive theory of multimedia learning is based on three assumptions: (a) visual

information and auditory information are processed by learners through two different channels, namely the visual channel and the auditory channel (dual-channel assumption); (b) each channel can process only a limited amount of information at a time (limited capacity assumption), and (c) learning occurs when learners engage in active cognitive processing of incoming information (active processing assumption).

Mayer's (2005) theory serves as the cornerstone of the Integrative Model of TPC (Schnoltz & Bannert, 2003). Although this model is based on similar assumptions, it differs from the cognitive theory of multimedia learning in two important respects. Schnoltz and Bannert's model (2003) differentiates between sensory modality and representational format. Sensory modality refers to sensory channels at the perceptual level – auditory channel (ears) and visual channel (eyes)--through which visual and auditory information enter working memory. Representational format, on the other hand, pertains to representational channels at the cognitive level --verbal channel and pictorial channel—through which mental models are constructed in a cognitive system. Unlike Mayer's (2005) theory that merges auditory---verbal and visual---pictorial channels, Schnoltz and Bannert's (2003) model suggests that visual information can be obtained not only through the visual modality, but also auditorily (e.g., from sound images). Furthermore, while Mayer's (2005) theory postulates the construction of two mental models—a verbal and a pictorial model—that are subsequently integrated into the cognitive system; the integrated model presupposes the construction of one mental model that combines information from both channels.

Of the three models, it is the contention of the authors of this study that The Integrative Model of TPC (Schnotz & Bannert, 2003) is the most appropriate for interpreting the findings of this study, hence this model will be taken as the theoretical framework for this study.

Relevant studies

The relationship between cognitive process and listening skills

Listening is an indispensable skill for language development. Although listening is regarded as a passive skill, listening comprehension is an active and fundamental language skill (Field, 2008). Further, listening and vision are inextricably linked and aural input is enhanced by visual information. Cooper's (1974) study results established a strong relationship between eye fixations on objects and the semantic representations of those items in aural input. Boland's (2004) findings further suggested that the probability of gazing and eye-fixation on the object increases when the object is mentioned in the narration.

In a more recent study by Cigerci and Gultekin (2017) on the use of digital stories to develop listening comprehension skills in primary school children, it was concluded that significant improvements in the listening comprehension skills were shown by the experimental group when provided with engaging listening activities during storytelling sessions. In an earlier experimental study conducted by Collen (2006), the children in the experimental group watched and listened to digital stories while the control group listened to the stories read aloud by the researcher. The listening activities were recorded, and the children were asked questions about what they heard, as well as the researcher, asked questions about the stories. Findings confirmed that children in the experimental group

were more focused and answered many questions correctly when quizzed on the story compared to the control group. A few studies echoed similar findings that support the importance of visuals accompanying narration in developing listening comprehension skills in young children (Abidin, et al. 2011, Verdugo & Belmonte, 2007). Whilst adequate research in understanding the cognitive process and reading comprehension in preliterate children is available, evidence is scarce on the use of visuals in the acquisition of listening comprehension abilities in young children.

Eye Movement & Comprehension Monitoring

Comprehension monitoring involves the process of analysing eye movements during reading activity to ascertain children's level of comprehension skills. Analysis of eye movements can provide data that informs the time course of reading and inferences about the underlying process that supports reading. Research has revealed that eye movements are a reliable source to predict children's comprehension level, even for young children of 2 years old (Annina et al., 2021; Koornneef & Mulders, 2017; Southwell et al., 2020). Findings of some studies further suggested that children were able to match words they heard with pictures (Kaefer et al., 2016; Marian et al., 2011; Takacs & Bus, 2018;).

Cognitive Processing/Visual information processing

Visual information processing during reading situations can be defined as the method of processing information seen and read and making sense of this information. Comprehension monitoring is therefore interrelated with the concept of visual information processing. Visual information processing involves examining and analysing eye movements specifically, time to first fixation, fixation duration, and fixation counts. Fixation duration and fixation counts have been widely recorded and collated in studies on visual information processing and its correlation to children's literacy and language proficiency.

Visual processing can be dependent on the condition of information presented during reading situations. The condition of information presented such as whether it contained visuals that accurately reflect the story or incongruent narration and/or text with pictures can affect children's visual processing hence their eye movements. Several studies have elucidated that for reading conditions that contained incongruent narration and pictures - the fixation duration was shown to be the longest and slightly longer than the congruent condition (Takacs & Bus, 2018; Su et al., 2018). This suggests that the incongruent reading conditions demand a higher cognitive load thus children generally fixate longer to process the mismatched text and pictures.

Characteristics of Story

In describing the characteristics of a story we are referring to whether the story has written text and pictures, pictures only, pictures with written text, or a story accompanied by audio narration. These characteristics have a correlation to children's literacy and language skills which include their ability to comprehend and ability to

comprehend and retell the story (Flack and Horst, 2017; Cigerci and Gutelking, 2017; Takacs and Bus, 2018).

Studies have shown that the impact of the picture on young children's comprehension and retelling of the story depends on a variety of circumstances (Flack & Horst 2017, Takacs & Bus, 2018; Wang, 2020). For example, Takacs and Bus (2018) found that children's retellings of stories were more accurate to the original narration when they heard the original narration with matching pictures as compared to all the other conditions such as the picture-only condition in which the children had to come up with their own story based on the given picture. Their study concluded that pictures alone are insufficient for children to retell the story accurately. The importance of key visuals in a story is also highlighted and reiterated in a more recent study. Wang (2020) asserted that preciseness of key visuals correlated with young children's understanding of picture books when it came to identifying and understanding the topic/theme of the book, attributes of the phenomena, and understanding characteristics and events of phenomena.

Multimodal characteristics of stories such as the combination of audio narration and key visuals can also facilitate listening skills (Cigerci & Gulteking, 2017). The multimodal aspects of audio narration and pictures were seen to heighten children's visual interest as shown in their fixation duration (Takacs & Bus, 2018). Most of the studies discussed so far were undertaken on children exposed to materials in their first/native language. According to Whitford and Joanisse (2018), bilingual children's simultaneous exposure to two reading systems could lead to different eye movement behavior compared to monolingual children, which may affect lexical information acquisition and assimilation in both languages. Therefore, understanding their cognitive processes will help teachers choose appropriate teaching approaches and materials for them.

Research Methodology

Participants

A total of 22 children aged between 5 to 6 were selected for this study. They were from a kindergarten in the Klang Valley. The kindergarten was selected out of availability. It was the only kindergarten that allowed the researchers to undertake the experiments amid the Covid-19 pandemic and it had only 22 children of the targeted age. The children were of Chinese origin and their L1 was Mandarin, however, they could speak and understand English but had not been exposed to any formal English reading programme, hence, they were classified as prereaders so English could be considered their other language (EAL). Permission from parents was obtained before the commencement of the research project.

Research Design

This study used an eye-tracking methodology to investigate the 22 prereaders' cognitive process when reading four short stories by examining the relationship between their eye movements and their story-telling ability.
They were exposed to four conditions:

(a) Congruent condition - oral narration, and a picture and a text that are congruent with the narration
(b) Incongruent condition – picture and text are the same as the congruent condition but the narration tells a different story
(c) Picture condition (Control) - only a picture with related text but no oral narration
(d) Text condition - only narration with related text

The study used a within-participant experimental design in which every child would participate in four conditions. All 24 possible narrative-picture matches were used across children. Children were evenly assigned to the different orders (as shown in Table 1).

Table 1
Assignment of children to the different conditions

Condition	Picture A	Picture B	Picture C	Picture D
1.Congruent picture + oral narration	Child A	Child B	Child C	Child D
2. Incongruent picture + oral narration	Child B	Child C	Child D	Child A
3. Only picture	Child C	Child D	Child A	Child B
4. Only written text + oral narration	Child D	Child A	Child B	Child C

Materials

For the congruent condition, four colored pictures were chosen for the study, and four narratives in English were constructed to describe the pictures. The content and language of each of the narratives were based on children's picture storybooks that were appropriate for children aged between 5 to 6 years old. For the incongruent condition, four other narratives, which had had similar main characters with each of the congruent pictures but d with different scenery, and images were constructed. These narratives were constructed using the same criteria as the narratives for the congruent condition, in terms of length (around 30 seconds), difficulty level, and cultural appropriacy. The first picture is provided in Appendix A together with accompanying materials for the four conditions. The second picture shows a mother waking her daughter up from bed in the morning. The third picture displays a girl in a hospital bed with a nurse attending to her, and the last picture illustrates a boy with his mother outside a toy shop. The narration for each story was read by a competent English teacher from Singapore.

Five visual elements were selected in each of the pictures as fixation points for the eye-tracking analysis, and the text was slightly modified to mention the objects in the text and narration. 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).

Research Procedures

Testing and intervention took place in a designated quiet location of the kindergarten with each session taking approximately 30 minutes per child.
The steps are as follows:

- First, the children had to go through the eye-tracking calibration process. Following that, a headphone was placed on each child and he/she was told to listen to the narration (if present) while looking at the monitor which had the eye tracker attached.
- Afterwards, the child was led upstairs to play a short game, such as building a Lego toy or completing a wooden puzzle. They were given 10 minutes to play the game which was included to ensure the children did not simply utter words of the story stored in their short-term memory without thinking.
- After the game, the child was asked to recall the story heard. For condition 3 (control) which had no narration, the child was asked to come up with a story based on what he/she saw. Each child went through the procedure four times.
- A Gazepoint GP3 HD eye tracker was used to collect eye movement data of the children and the data were analysed statistically. Areas of interest include the image (that is the picture as a whole), text (the text as a whole), the objects (the selected five visual elements as seen in the picture), and keywords (the five visual elements seen as words in the text).
- Retellings were prompted by general questions such as "Can you tell me more? What else did you hear?" (when necessary). Children were allowed to retell their stories in English or in Mandarin. Children's retellings were transcribed and coded. A rating scale was devised and this was used to guide the three independent raters invited to undertake the ratings. The raters were academics with at least a Master's degree in Education/Early Childhood Education. A meeting was held for the raters to discuss and compare the scores for each child. The final score was based on a consensus among the three raters. The scale scored the storytelling on a five-point scale and the scores for all conditions were added and based on them, each prereader's storytelling performance (STP) was then categorisation as shown below:

Table 1
Categorisation of storytelling performance (STP)

Rating scale	Category
0 mark = with no recollection beyond what is obvious from the picture.	Weak
1 mark =With some recollection of narration beyond recollection from the picture.	Weak
2 marks = With about half recollection of narration beyond recollection from the picture.	Average
3 marks= Recollect most of the events in the story beyond recollection from the picture.	Good
4 marks= Summarising all the main events very well in the story beyond recollection from the picture.	Good

- The findings of the STP were then matched with those from their eye movements.

Apparatus used

A Gazepoint GP3 HD eye tracker was used to collect eye movement data. The sampling rate of this eye tracker is 150 Hz and the angle error is 0.5 to 1 degree. It has two infrared light-emitting diodes (LEDs) that provide light that would reflect on the cornea of the left eye to detect eye movement. The eye tracker is able to work at an optimal distance of 65 cm away from the eye. The Gazepoint GP3 HD eye-tracking system includes two software tools: Gazepoint Analysis and Gazepoint control. These tools were used to define the Areas of Interest (AOI) and generate metrices. The metric used for this particular study was Proportion Fixation Duration(PFD)(i.e, percentage of fixation time within an AOI. The eye movement data were collected at a 150 Hz sampling rate. Participants' head movements were minimised using a standard chinrest 60 cm from the monitor.

The eye tracker and its accompanying software recorded the position each participant was looking at on the screen 150 times each second. A calibration process was undertaken at the beginning where participants were asked to look at a series of dots on the screen. By analysing the position of the eyes of the participant, the software was capable of using the position of the eye to calculate where on the screen the participant was looking. These data were then saved in the form of coordinates.

The Gazepoint Analysis software also recorded what was on the screen at the time. By combining the coordinated information and the recording, Gazepoint Analysis was able to recreate the screen including the point on the screen the participant was looking at. Gazepoint Analysis was also used to generate reports on how often certain parts of the screen were looked at through the usage of Dynamic Areas of Interest. This allowed the areas of interest for different images and scenarios to be set up and defined. The statistics from Gazepoint Analysis were analysed for this study.

Data Analysis Procedures

Two types of data analysis were performed on the eye-tracking data for this study. A descriptive analysis of Proportional Fixation Duration (PFD) on the four predefined areas of interest – Image, Text, Object, and Keyword – was undertaken. Inferential statistics, in the form of comparison of mean scores using mixed factorial ANOVAs, with storytelling performance [good, average, or weak] and gender as the between-subject variables (whereas AOI and conditions as the within-subject variables). Post-hoc analyses were also performed using IBM SPSS statistical software.

AOI for the whole picture and the five targets were made before the start of the experiment so that the software could identify these areas and elements for data analysis purposes. As a set, these comparisons are not mutually orthogonal, but the fact that each comparison tested for a part of the expected pattern overrides the need for orthogonality (Keppel & Wickens, 2004). In addition, the stories told by the prereaders were analysed qualitatively.

Research Finding

Quantitative analysis of eye-tracking data under the four conditions

Sue et al. (in press) reported in an earlier study eye-tracking patterns similar to that of Verhallen and Bus (2011), Montag et al. (2015), and Takacs and Bus (2018). The results revealed that pictures positively affected the understanding of narration of prereaders with prereaders paying less attention to the story when only a text was present. It was further found that the fixation time of the predefined AOI (Areas of Interest i.e. image, text, objects, and keywords) was significantly higher in the congruent condition than in the control condition. The findings also demonstrated that the prereaders were not unduly distracted by a non-matching picture in the incongruent condition.

Quantitative analysis of the interactions between proportional fixation time, gender, and story-telling performance (STP)

For this study, inferential statistics were used to analyse the interactions between PFD, gender, and story-telling performance. Mauchly's test of significance indicated that the test was significant ($p < .001$) for both conditions and combined AOI. As the assumption of sphericity was not met, the Greenhouse-Geisser correction was applied. The repeated-measures ANOVA with a Greenhouse-Geisser correction showed that there was no significant difference between conditions and story-telling performance [$F(3.57, 23.23) = 1.740, p = .180, \eta_p^2 = .16$]. Similarly, there was also no interaction between conditions, gender and story-telling performance [$F(3.57, 23.23) = 0.734, p = .564, \eta_p^2 = .07$]; AOI and story-telling performance [$F(2.20, 14.32) = 0.681, p = .535, \eta_p^2 = .11$]; and AOI, gender and story-telling performance [$F(2.20, 14.32) = 2.724, p = .096, \eta_p^2 = .14$].

The findings clearly showed that the prereaders' story-telling performance did not depend on treatment conditions, gender, or proportional fixation time on the combined AOI. These findings were very different from earlier studies by Flack and Horst (2017), Takacs and Bus (2018), and Wang (2020).

A review of findings from 4.1 and 4.2 showed two very contrasting patterns. The finding of the current research project matched that of earlier studies with respect to treatment conditions and eye-tracking fixation patterns, but unlike the earlier studies, they showed that the prereaders' STP was not significantly different between gender and the four different conditions. It would be intriguing to explore the underlying reasons for this phenomenon. None of the abovementioned studies attempted to explore prereaders' STP by analysing the prereaders' stories qualitatively. Hence, in doing so, the current study would be contributing new knowledge to this field.

Qualitative analysis of the relationship between fixation time, gender, and story-telling performance

Table 2
Female prereaders' storytelling performance according to treatment conditions

Prereader (Girl)		Story-telling performance (STP) in different conditions			
	Overall STP	Congruent (narration with matching picture)	Incongruent (narration with non-matching picture)	Picture (picture with the test without narration)	Text (narration with text without picture)
G1	AA	Good	Average	Average	Weak
G2	AA	Average	Average	Weak	Average
G3	A	Average	Average	Weak	Good
G4	BA	Weak	Average	Weak	Weak
G5	AA	Average	Average	Good	Weak
G6	A	Average	Weak	Good	Weak
G7	P	Weak	Weak	Weak	Weak
G8	A	Good	Weak	Average	Weak
G9	P	Weak	Weak	Weak	Weak

KEY
Ranking of overall STP: *Ranking of STP*:
Weak STP in one condition = Above Average (AA)   0 &1 = weak
Weak STP in two conditions = Average (A)   2 = average
Weak STP in three conditions= Below Average (BA)   3 & 4 = Good
Weak STP in four conditions= Poor (P)

A descriptive comparison revealed that the highest number of weak STP came from text condition (7 out of 9 weak cases), followed by picture condition (5 out of 9 weak cases), and incongruent condition (4 out of 9 weak cases), with the best STP coming from the congruent condition (only 3 weak cases). Although these descriptive findings borne similarity to findings of earlier studies (e.g. Takacs & Bus, 2018) they could not be taken as proof of similarity in results as the findings from inferential statistics revealed no significant difference between story-telling performance (STP) and treatment conditions in the case of female prereaders. To have a better understanding of this phenomenon, the researchers decided to undertake a qualitative analysis of the stories told to gain deeper insights into the situation. This was done by undertaking a thematic analysis of each child's STP for each of the conditions.

Analysis of the storytelling transcripts revealed that female prereaders' failure to perform well in the **Text condition** included inability to remember (G1, G5, G8) which they mentioned specifically, not paying attention, and repeating a story they heard earlier (G4, G6, G9) or presenting a totally unrelated story (G7). These patterns cut across different STP as shown in Table 2. This suggests that even prereaders with good STP scores could also get distracted when presented with only text and narration.

The main weakness of the female prereaders with regard to the **Picture condition** was the inability to describe a story well without hearing the narration and that was evident even in prereaders better in storytelling (G2, G3). In the case of prereaders with poor STP scores (G4, G7, and G9), they also produced utterances that were disjointed or incomplete.

In the case of the **Incongruent condition**, the reasons for their weaknesses varied. G6 could describe the picture well but she could not remember anything regarding the narration. It was noted she was also excellent in the picture condition suggesting that she had the tendency to focus more on pictures than narration. As for G7, her story had nothing to do with the picture or the narration, whereas G8 and G9 came up with stories that were a mixture of ideas from the picture and narration which showed that they were confused.

Only three prereaders (two poor and one below average) performed badly in the **Congruent condition**. They appeared to be suffering from attention deficit. G7 and G9 mostly created their own stories and paid very little attention to the narration and the picture for all four conditions. G4 to a large extent did the same thing although she did pay some attention to the incongruent condition.

Table 3
Male prereaders' storytelling performance according to treatment conditions

Child (Boy)		Story-telling performance (STP) in different conditions			
	Overall STP	Congruent (narration with matching picture)	Incongruent (narration with non-matching picture)	Picture (picture with test without narration)	Text (narration with text without picture)
B1	AA	Average	Average	Weak	Average
B2	AA	Average	Average	Average	Weak
B3	P	Weak	Weak	Weak	Weak
B4	A	Average	Weak	Average	Weak
B5	P	Weak	Weak	Weak	Weak
B6	A	Average	Average	Weak	Weak
B7	AA	Good	Weak	Average	Good
B8	BA	Weak	Average	Weak	Weak
B9	P	Weak	Weak	Weak	Weak
B10	P	Weak	Weak	Weak	Weak

KEY
Ranking of overall STP: *Ranking of STP*:
Weak STP in one condition = Above Average (AA)   0 &1 = weak
Weak STP in two conditions = Average (A)   2 = average
Weak STP in three conditions= Below Average (BA)   3 & 4 = Good
Weak STP in four conditions= Poor (P)

The descriptive comparison for the male prereaders also indicated the highest number of weak STP came from text condition (8 out of 10 cases), followed closely by the picture condition (7 out of 10 weak cases), and incongruent condition (6 out of 10 weak cases) and the least number of weak cases was the congruent condition (5 out of 10 cases).

The qualitative analysis of the stories revealed that for the **Text condition**, the male prereaders could only come up with one to two short sentences with a few words related to the narration they heard (B2, B3, B5) that did not make much sense or short confused stories (B6, B8, B10). One even said he could not remember anything (B9).

The male prereaders were also weaker than the female prereaders with regard to the **Picture condition**. Three came up with one or two brief, not very meaningful sentences with a few words related to the picture they viewed (B1, B8, B9), three with two to three totally unrelated sentences (B3, B5, B6) and one could not say anything (B10).

For the **Incongruent condition**, the male prereaders were also weaker than their female counterparts. Two could not say anything (B5, B10), one gave one short, not a very meaningful sentence with three words related to the picture they saw (B3), and three gave a short mixed-up story of the narration and the picture (B4, B7, B9).

Similar to the female prereaders, only poor and below-average children performed badly in the **Congruent condition**. The very weak student could not say anything (B10), two gave one brief, not meaningful sentence(s) with three words related to the narration they heard (B3, B5), and two clearly were not paying attention: one repeated an earlier story he heard (B8) and another managed to pick up three words from the narrative and added that to his own story (B9).

The quantitative findings of this study revealed that there were no significant differences between conditions, gender, and STP which showed that prereaders' STP did not depend on treatment conditions, gender, or average fixation time on the combined AOI.

Discussion of Findings

The quantitative findings of this study revealed that there were no significant differences between conditions, gender, and STP which showed that prereaders' STP did not depend on treatment conditions, gender, or average fixation time on the combined AOI. These findings were very different from earlier studies by Flack and Horst (2017), Takacs and Bus (2018), and Wang (2020).

The qualitative analysis of failure to perform well in story-telling in the Text condition suggests that the lack of pictures despite the presence of narration can cause a lack of attention and inability to remember, especially in the case of the prereaders with weak story-telling abilities. Thus, in such cases, even though the prereaders might have appeared to be looking at text, they were not absorbing much of what they saw or heard. This could be due to the fact that the text-only condition provided limited possibilities for children to create mental representations/images of the narrated story leading to inhibition of mental processing and poor retrieval of the narration. There is also the possibility that they were distracted by the written text and hence did not focus on the narration because they were prereaders (still learning to read).

The main problem in the case of the Picture condition was the prereaders' inability to describe a story without hearing the narration and that was evident even in prereaders better in storytelling. Schnoltz and Bannert's integrated model (2003) presupposes that sensory information entering the working memory from the visual modality and auditorily modality will be combined as one mental model. This implies that when information from one modality is missing it will affect the children's abilities in remembering and narrating a story. In the case of the Picture condition, the children were only exposed to a picture and they had to draw on their memory of the picture to tell a

story. In doing so weaker children would be further handicapped by their poor storytelling abilities.

Schnoltz and Bannert's model can also be drawn upon to explain the situation in the case of weak prereaders for the Incongruent condition. Due to the mismatch between picture and narration, some of them were distracted and confused leading to them not being to come up with an integrated model that would combine information from both channels efficiently. Male children seemed more confused and less willing to try. This failure could be attributed to split attention and a higher level of cognitive load in attending simultaneously to a narration that is incongruent with the picture. In the case of the Congruent condition, generally, STP was better and this could be attributed to the match in the information for the two modalities.

Since no other studies have used the qualitative approach to analyse the STP of prereaders in the same way as this study has done, it is not possible to draw from past studies. However, it is possible to deduce that affective factors such as the state of mind of the participants, their abilities to concentrate, and their level of attentiveness might have come into play in influencing the reliability and validity of eye-tracking data of these young prereaders, especially since they were exposed to materials, not in their native/first language. Past studies done in the 1990s to 2000s have looked into pre-schoolers' attention and engagement and have concluded that they displayed a lower level of attentiveness and engagement when compared to school-going children. For example, research using vigilance tasks such as the Continuous Performance Test (CPT) has shown that pre-schoolers might be expected to make more omission errors than older children, have greater difficulty maintaining the task instructions in working memory, and have related difficulties with response inhibition, particularly in the presence of distractors. They also have more difficulty maintaining their attention over a long period of time (Akshoomoff, 2002; Corkum et al., 1995; Levy, 1980; Prather et al. 1995). All of these could have led to the lack of significance in the quantitative findings of this study. In addition to that their level of language proficiency in English, and their general ability in communicating verbally are also important factors to take into consideration. Other eye-tracking studies done on English as Another Language (EAL) were done mostly on school-going children reading different types of texts (e.g. Connor et al., 2015; Hessel et al., 2020; Kim et al., 2018). Very few were done on pre-schoolers exposed to another language (except for Verhallen & Bus 2010, 2011). More importantly, none of these studies explored the issue of attentiveness and engagement. Hence it is not possible to compare their results with those of the current study. This suggests that further investigation needs to be undertaken to shed more light on these issues in the EAL context.

Implications and Conclusion

Based on the eye-tracking data, it can be concluded that having any picture positively affected the understanding of the narration of the prereaders. Thus, it appears that they were more focused on the presence of an image. Of the four conditions. the congruent condition was the most effective at building this understanding. The Incongruent condition and the Control condition, where there was no narration, were next in effectiveness. The condition with text and narration only was the least effective at positively affecting the understanding of narration.

The eye-tracking results further suggested that prereaders did use visual attention to build understanding. It appeared that even without narration or with a picture that did not match the narration, the eye movements showed improved attention in the presence of an image. These findings suggest that images are important and reading materials designed for prereaders should have them to improve focus and attention to the story.

However, further analysis of the eye-tracking data revealed that the prereaders' story-telling performance did not depend on treatment conditions, gender, or eye fixation on the combined AOI. These findings differed from earlier studies. The qualitative analysis suggested affective factors such as the state of mind of the participants, their abilities to concentrate, and their level of attentiveness could have affected the prereaders' STP. Hence they may appear to be focusing on the screen but their attention might have been somewhere else or they could have been confused Duchowski (2017). This is referred to as overt attention. Thus, assuming that the visual attention of a subject is associated with their point of fixation may not be correct in some cases as shown in this study. Thus, other methods of identifying overt attention need to be considered. The current study is an experiment in that direction. It could also be inferred that level of language proficiency and their general ability in communicating verbally could have been contributory factors (An et al., 2017).

While eye movements differed between the experimental conditions, they did not seem to consistently predict higher-order storytelling performance. Individual differences and gender differences in reading ability might have dampened the link between eye movement and group-level comprehension performance (Zhan et al., 2020). Although the current study found no strong causal relationship between fixation patterns and storytelling performance, our work is a step toward developing integrative theories of story comprehension by systematically measuring overt attention. Future research should utilise eye-tracking technology to elucidate how differences in story comprehension arise during reading, and to what extent looking strategies are associated with children's gender and language backgrounds.

It is further proposed that one of the reasons for the differences between the findings found in earlier studies and this study could possibly be due to a lack of language proficiency in the language used in the eye-tracking experiments. In the earlier studies, the pre-schoolers were exposed mostly to their first/native language whereas in this case they were exposed to another Language. This could have affected some children's ability in comprehending and narrating the stories, leading to confusion and a lack of attention. This is an area that needs to be investigated in future studies. In addition to that, there is a need to look into "attentiveness and active engagement" using reliable research methods to find out to what extent they affect the prereaders' STP. This will enable a more comprehensive understanding of the cognitive processes of these prereaders. These could be considered limitations of the study as these variables were not taken into consideration. In addition to that, the sample size of this study was limited to only 22 children from one kindergarten. The study needs to be extended to a larger sample size and also include children from other ethnic origins to find whether the findings are applicable to them.

Given that recent neuroimaging studies have detected electrophysiological markers – greater involvement of occipital brain regions – supporting story comprehension in pre-schoolers using encephalography (EEG), future research could apply this advanced technique to identify the relationship between cortical connectivity

and visual attention during narrative comprehension tasks, and to what extent these neurophysiological measures predict vocabulary growth in young children.

Funding

References

Abidin, M. J. Z., Pour-Mohammadi, M., Souriyavongsa, T., Da, C. & Ong, L. K. (2011). Improving listening comprehension among Malay preschool children using digital stories. *International Journal of Humanities and Social Science*, 1(14), 159-164. http://www.ijhssnet.com/journals/Vol_1_No_14_October_2011/21.pdf

Akshoomoff, Natacha (2002). Selective attention and active engagement in young children. *Developmental Neuropsychology*, 22(3), 625–642. https://doi.org/10.1207/S15326942DN2203_4

An L., Wang Y., & Sun Y. (2017) Reading words or pictures: Eye movement patterns in adults and children differ by age group and Receptive Language ability. *Front. Psychol.* 8:791. https//:doi: 10.3389/fpsyg.2017.00791.

Annina K. H., Nation, K. & Victoria, A. M. (2021) Comprehension monitoring during reading: An eye-tracking study with children learning English as an additional language, *Scientific Studies of Reading*, 25:2, 159-178. https//doi:10.1080/10888438.2020.1740227

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & J. C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 51–76). London, UK: Psychology Press

Cigerci, F. M., & Gultekin, M. (2017). Use of digital stories to develop listening comprehension skills. *Issues in Educational Research*, 27(2), 252-268. http://www.iier.org.au/iier27/cigerci.pdf

Collen, L. (2006). The digital and traditional storytimes research project: Using digitized books for preschool group storytimes. *Children and Libraries*, 4(3), 8-18. https://worldlibraries.dom.edu/index.php/worldlib/article/view/43/68

Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L., & Morrison, F. J. (20015) Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: An eye-movement study. Scientific Studies of Reading, 19 (2), 114–134. https://doi.org/10.1080/10888438.2014.943905

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 84107. https://doi.org/10.1016/0010-0285(74)90005-X

Corkum, V., Byrne, J. M., & Ellsworth, C. (1995). Clinical assessment of sustained attention preschoolers. *Child Neuropsychology,* 1, 3–18. https://doi.org/10.1080/09297049508401338

Evans, M. A. & Saint-Aubin, J. (2005). What children are looking during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science*, 16, 913-920. https://doi.org/10.1111/j.1467-9280.2005.01636.x

Field, J. (2008). *Listening in the language classroom*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511575945

Flack, Z. M., & Horst, J. S. (2017). Two sides to every story: Children learn words better from one story page at a time. *Infant and Child Development*. Advance Online Publication. https://doi.org/10.1002/icd.2047

Gong, Z., Levy, B. A. (2009). Four year old children's acquisition of print knowledge during electronic storybook reading. *Reading and Writing*, 22, 889–905. https://doi.10.1007/s11145-008-9130-1

Hessel, A. K., Nation, K. & Murphy, V. A. (2020) Comprehension monitoring during reading: an eye-tracking study with children learning English as an additional language. *Scientific Studies of Reading*. 25(2). https://doi.org/10.1080/10888438. 2020.1740227

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, 229-243. https://doi.org/10.1111/j.1467-9817.2005.00267.x

Justice, L. M., Pullen, P. C., Pence, K. (2008). Influence of verbal and nonverbal references to print on pre-schoolers' visual attention to print during storybook reading. *Developmental Psychology*, 44, 855–866. https://doi.org/10.1037/0012-1649.44.3. 855

Kaefer, T., Pinkham, A. M., & Neuman, S. B. (2016) Seeing and knowing: Attention to illustrations during storybook reading and narrative comprehension in 2-year-olds. *Infant and Child Development*, 26(5). https://doi.org/10.1002/icd.2018

Keppel, G., & Wickens, T. D. (2004). *Design and analysis*: A researcher's handbook (4th ed.).Upper Saddle River, NJ: Pearson Prentice Hall.

Kim, Y. S., Vorstius, C., & Radach, R. (2018). Does online comprehension monitoring make a unique contribution to reading comprehension in beginning readers? Evidence from eye movements. *Scientific Studies of Reading*, 22(5), 367–383. doi:10.1080/10888438.2018.1457680

Koornneef, A., & Mulders, I. (2017). Can we 'read' the eye-movement patterns of readers? Unraveling the relationship between reading profiles and processing strategies. *Journal of Psycholinguistic Research*, 46(1), 39–56. https://doi.org/10.1007/s10936-016-9418-2

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(1), 103703. https://doi.org/10.1016/j. compedu.2019.103703

Levy, F. (1980). The development of sustained attention (vigilance) and inhibition in children: Some normative data. *Journal of Child Psychology and Psychiatry, 21,* 77–84. https://doi.org/10.1111/j.1469-7610.1980.tb00018.x

Lin, D., Chen, G., Liu, Y., Liu, J., Pan, J., & Mo, L. (2018). Tracking the eye movement of four years old children learning Chinese words. *Journal of Psycholinguistic Research*, 47(1), 79–93. https://doi.org/10.1007/s10936-017-9515-x

Liu B. G. (2011) *Eye movement study on the development of Chinese character consciousness in picture book reading for children aged 4–6*. Shanghai: East China Normal University.

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. https://doi.org/10.1177/0956797615594361

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Ochs, E., & Schieffelin, B. (2009). Language acquisition and socialization: Three developmental stories and their implications. In A. Duranti (Ed.), *Linguistic anthropology: A reader* (2nd ed.) (pp. 296–328). Malden, MA: Wiley-Blackwell

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Paivio, A. (1971). *Imagery and Verbal processes*. New York: Holt, Rinehart, & Winston.

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

Prather, P. A., Sarmento, N., & Alexander, A. (1995). Development of vigilance in preschoolers. *Journal of the International Neuropsychological Society,* 1, 153. https://doi.org/10.1017/S1355617700001843

Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction,* 13,141-156. https://doi.org/10.1016/S0959-4752%2802%2900017-8

Southwell, R., Gregg, J., Bixler, R. & D'Mello, S. K. (2020), What eye movements reveal about later comprehension of long connected texts. *Cogn Sci*, 44: e12905. https://doi.org/10.1111/cogs.12905

Su, Z., Wang, Y., Sun, Y., Ding, J. & Ma, Z. (2018) Reading independently and reading with a narrator: Eye movement patterns of children with different receptive vocabularies. *Frontiers in Psychology.* 24 September 2018. https://doi.org/10.3389/fpsyg.2018.01753

Sue, C. H., Thang, S. M., Wong, H. K., Tan, P. S., Loo F.L. & Ahju, R. (in press). 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 & Teaching (IJCALLT).*

Takacs, Z. K., & Bus, A. G. (2016). Benefit of motion in animated storybooks for children's visual attention and story comprehension. An eye-tracking study. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2016.01591

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. https://doi.org/10.1016/j.jecp.2018.04.013

Tanenhaus, 2007. Eye movements and spoken language processing. In R.P.G. van Gompel, M. H. Fisher, W. S. Murray & R. L. Hill (Eds). *Eye movements: A window on mind and brain* (pp.443-469). Oxford, UK: Elsevier.

Verdugo, D. R. & Belmonte, I. A. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English Language. *Learning & Technology*, 11(1), 87-101. https://www.lltjournal.org/item/10125-44090/

Verhallen, M. J. & Bus, A. G. (2010). Low-Income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology*, 102(1), 54-61. https://doi.org/10.1037/A0017133

Verhallen, M. J. & Bus, A. G. (2011). Young second language learners' visual attention to illustrations in storybooks. *Journal of Early Childhood Literacy,* 11, 480-500. https://doi.org/10.1177%2F1468798411416785

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. https://doi.org/10.1177%2F2513850219888029

Whitford, V., & Joanisse, M. F. (2018). Do eye movements reveal differences between monolingual and bilingual children's first-language and second-language reading? A focus on word frequency effects. *Journal of Experimental Child Psychology*, 173, 318-337. https://doi.org/10.1016/j.jecp.2018.03.014

Zhan, Z., Wu, J., Mei, H., Wu, Q. & Fong, P. S. W. (2020). Individual difference on reading ability tested by eye-tracking: From perspective of gender. *Interactive Technology and Smart Education*, 17 (3), 267-283. https://doi.org/10.1108/ITSE-12-2019-0082

Appendix A

Materials for Picture A

Condition	On screen	Oral narration
Congruent	 It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.	It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.
Incongruent		John is helping Mama in the kitchen. Mama wants to cook green bean soup for John. There is a bag of green beans, a cup of water, and a big bowl on the table. "Can

		I help you pour the green beans into the bowl?" John asks. "Yes, you can but be careful not to spill any on the table". Mama says. Mama laughs when John accidentally pours some onto her hands. Some fall on the floor too.
	It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.	
Picture Control		No oral narration
	It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.	
	No picture It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go	It is a windy day. Mama and John are sitting in the back garden. They are eating a bowl of green bean porridge. "Look at the wind blowing the shirts and pants around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her pink and white shopping bag by her side. She plans to go

	shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.	shopping with John after this. They are having a good time, enjoying the breeze and looking at the cabbage patch.

The visual elements selected as fixation points are underlined.

It is a windy day. Mama and John are sitting in the back garden. They are eating **a bowl of green bean porridge**. "Look at the wind blowing the **shirts** and **pants** around", mama says. Yes, they look like people flying around" John shouts happily. Mama has her **pink and white shopping bag** by her side. She plans to go shopping with John after this. They are having a good time, enjoying the breeze and looking at the **cabbage patch**.