Learning and Instruction 84 (2023) 101731

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Do easy-to-read adaptations really facilitate sentence processing for adults with a lower level of education? An experimental eye-tracking study

Miriam Rivero-Contreras [a,*], Paul E. Engelhardt [b], David Saldaña [a]

[a] *Universidad de Sevilla, Sevilla, Spain*
[b] *University of East Anglia, Norwich, UK*

A R T I C L E I N F O

Keywords:
Visual support
Lexical simplification
Low level of education
Sentence processing
Eye movements

A B S T R A C T

The *Easy-to-Read* guidelines recommend visual support and lexical simplification to facilitate text processing, but few studies have empirically verified the efficacy of these guidelines. This study examined the influence of these recommendations on sentence processing by examining eye movements at the text- and word-level in adult readers. We tested 30 non-university adults (low education level) and 30 university adults (high education level). The experimental task consisted of 60 sentences. Half were accompanied by an image and half were not, and half contained a low-frequency word and half a high-frequency word. Results showed that visual support and lexical simplification facilitated processing in both groups of adults, and non-university adults were significantly slower than university adults at sentence processing. However, lexical simplification resulted in faster processing in the non-university adults' group. Conclusions focus on the mechanisms in which both adaptations benefit readers, and practical implications for reading comprehension.

1. Introduction

Reading is an extremely complex cognitive activity, which not only involves decoding words, but ultimately aims at an understanding of what has been read and a transformation of the reader's prior knowledge. According to the latest report from the Program for the International Assessment of Adult Competencies (PIAAC), the average scores obtained in reading comprehension by adults between 16 and 65 years of age across all participating countries was Level 2 (Ministerio de Educación, Cultura y Deporte, 2013). This level indicates that the readers can complete simple reading tasks and that they can locate information in a short text. However, they will have difficulty extracting information from longer and more complex texts. Perhaps more surprisingly, many developed countries fall well below this average, including the United States and European countries such as Spain and Italy.

In a meta-analysis by Tighe and Schatschneider (2016b) with 16 independent studies and 2707 participants, it was observed that six reading skills were highly related to reading comprehension of struggling adult readers, one of them being (oral) vocabulary (*Hedges' g* = 0.52). In addition, research has indicated that when other reading skills are controlled for, such as decoding (Hall et al., 2014; Taylor et al., 2012), fluency (Taylor et al., 2012), and morphological awareness (Fracasso et al., 2014; Tighe & Schatschneider, 2016a), vocabulary remains a significant predictor of comprehension, particularly in struggling adult readers, with medium and small effect sizes (F-Squared, $f^2 =$ 0.20 for Hall et al., 2014; 0.14 for Taylor et al., 2012; 0.10 for Fracasso et al., 2014; 0.05 for Tighe & Schatschneider, 2016a). Therefore, vocabulary seems to play a key important role in reading comprehension.

Given the importance of vocabulary, the level of lexical complexity is a key consideration when making reading materials accessible for those who have difficulties in reading comprehension or those with low literacy levels. The International Federation of Library Associations and Institutions (IFLA) and the ILSMH European Association, known today as Inclusion Europe, have proposed general recommendations and guidelines to make reading materials more accessible (Freyhoff et al., 1998; Tronbacke, 1997). The materials developed within these guidelines and recommendations are called *"Easy-to-Read"*. Likewise, in Spain, guidelines and recommendations for the preparation of documents in Easy-to-Read style have been published (Asociación Española de Normalización, 2018). Two important recommendations contained within the guidelines relate to the use of simple language with frequently used words and to the use of visual support to accompany text, both of which are the focus of this paper. However, since perceived

* Corresponding author.
E-mail address: mrivero5@us.es (M. Rivero-Contreras).

https://doi.org/10.1016/j.learninstruc.2022.101731
Received 31 March 2021; Received in revised form 13 December 2022; Accepted 19 December 2022
Available online 28 December 2022

text difficulty is the main recommended measure for testing successful adaptations, little attention has been paid in that field to the existing empirical literature on the specific effects of visual support and lexical simplification on reading comprehension and processing (Asociación Española de Normalización, 2018).

1.1. Visual support and lexical simplification

Visual support (i.e., visual representation) has a prominent position in Easy-to-Read guidelines. Images, illustrations, and/or symbols can accompany text, so that if a reader encounters an unknown word, the visual representation can be used to bootstrap understanding of the text, allowing the reader to obtain the intended meaning. It is well known that children partially learn to read using various kinds of visual support. It has been proposed that illustrations can provide information about content that is not part of everyday experience (Hibbing & Rankin-Erickson, 2003), and can help when text contains unfamiliar words and/or complex grammar (Montag et al., 2015). In adults, it might seem that images are no longer relevant for understanding written information, since books for adult and skilled readers usually only contain text or very few visuals.

However, some research suggests that images continue to facilitate lexical access especially to familiar words, with a large effect size ($g = -1.02$, Qu et al., 2016). In addition, several studies with individuals with comprehension difficulties and/or disability have directly evaluated the facilitation of comprehension in easy and difficult texts with the use of images or symbols, although with some conflicting results. For example, Dye et al. (2007) evaluated the ability of participants with intellectual disabilities to give consent with a questionnaire, accompanied by photographs that complemented the information, compared to a control version (information followed by the questionnaire) and a section version (information presented in sections and asking questions after each). They found no significant differences in the ability to consent between the experimental conditions.

Poncelas and Murphy (2007) and Jones et al. (2007) examined the influence of texts with symbols on the reading comprehension of people with intellectual disabilities and people with learning disabilities, respectively. While Poncelas and Murphy found relatively low levels of comprehension, with no significant differences between text only and text with symbols, Jones et al. (2007) did find an impact of visuals. Therefore, the type of visual support that accompanies the text seems to be important, and it also seems that any type of visual support does not serve to support the reading comprehension of every population with disabilities. Also, the effect sizes of these studies were mostly small (*Hedges' g* $= -0.26$ for Dye et al., 2007, p. 0.22 for Jones et al., 2007, p. 0.49 for Poncelas & Murphy, 2007).

However, a recent meta-analysis examining 39 experimental studies (20 with college students or older adults) found that the inclusion of graphics had a moderate overall positive effect (*Hedges' g* $= 0.39$) on students' reading comprehension, regardless grade level (elementary, secondary, and adults), suggesting that visuals can improve reading comprehension for students without disabilities or specific comprehension issues (Guo et al., 2020).

Lexical simplification is another aspect considered in Easy-to-Read guidelines. This consists of the substitution of low-frequency words, which may be unfamiliar to many people, with higher frequency words, which are more commonly known and more often used in oral language. Lexical simplification is considered in these guidelines because vocabulary is known to be a significant predictor of reading comprehension, especially in adult struggling readers (Tighe & Schatschneider, 2016b) and in less proficient adult readers who do not use phonological codes for word recognition (Chace et al., 2005) or do so less efficiently (Binder & Borecki, 2008). Word frequency is a (lexical) variable that has been extensively studied in the field of psycholinguistics and has a rich history in the eye-movement reading literature (e.g., Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986; Staub et al., 2010). Word frequency is correlated with the speed of lexical access, such that higher-frequency words are processed faster than low-frequency words (e.g., Ashby et al., 2005; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner & Fischer, 1996; Rayner et al., 1996). Estimates suggest less frequent words receive 20–60 ms longer fixation durations, even after controlling for word length.

1.2. Eye-movements measures

Eye tracking has been used to investigate the cognitive processes of young readers while reading illustrated texts (Hannus & Hyönä, 1999; Jian, 2017; Mason et al., 2015; Mason, Pluchino, et al., 2013; Mason, Tornatora, & Pluchino, 2013). Previous studies using eye-tracking measures support the idea that the presence of illustrations improves reading comprehension and learning of the illustrated text content, although most suggest that it only occurs in those with high abilities. For example, Hannus and Hyönä (1999) studied the effects of illustrations on the learning of authentic textbook materials among 10-year-old primary school children with high and low intellectual ability. They found that comprehension scores improved with the presence of illustrations for high-ability children, but not for low-ability children. Moreover, although no differences were found in the amount of time spent inspecting textbook illustrations, different patterns were observed: high-ability children made more back-and-forth looking between text and a corresponding illustration. Along the same lines, Mason et al. (2013b) evaluated the text and graphics processing of fourth grade students while reading an illustrated science text and found that those students with higher performance on learning tasks were associated with higher integrative text-graphics processing (a longer inspection time of the picture during the first encounter with it, more integrative transitions from one representation to the other, longer fixation time on the illustration while re-reading the text, and a longer fixation time on the text while re-inspecting the picture). Likewise, Jian (2017) evaluated the reading strategies and comprehension of illustrated biology texts from an elementary school science textbook in two groups with different visual literacy: fourth graders and university students. The results showed a better performance of the university students, in addition to different eye movement patterns. Adults exhibited two-way reading paths for both text and illustrations, whereas fourth graders' eye fixations only went back and forth within paragraphs of text and between illustrations. Therefore, research studies in this area support the idea that the greater the ability to integrate both elements, text and visual support, the greater reading comprehension and learning.

Eye tracking also permits the analysis of cognitive processing related to reading comprehension (for a review, see Rayner, 1998). A large body of evidence suggests that some variables are related to both lexical and higher-level linguistic processing, and that these variables affect the duration of fixations. There are also studies which suggest that reading ability may influence word processing time and re-reading patterns, such that unskilled readers (or readers with language/learning disabilities) will have longer fixation durations and an increased number of regressions (Everatt & Underwood, 1994; Rayner, 1998; Underwood et al., 1990). Ashby et al. (2005) compared the reading of 'highly skilled' and 'average proficiency' readers, who were grouped according to performance on vocabulary and comprehension tests. They observed that average-proficiency readers read target words more slowly, made more regressions, and spent more time re-reading than highly skilled readers, even in relatively simple sentences. Similarly, Barnes et al. (2017) investigated the reading skill of adults enrolled in a basic education program (i.e., a low literacy sample). Participants showed longer fixation durations, made a greater number of regressions, and showed longer overall reading times, than did "normatively" educated adults, with large effect sizes (g between 1.37 and 2.10). According to Barnes et al., this pattern of results was due to lower reading fluency, poor ability to recognize and read words, and poor comprehension. However, in this study participants read passages aloud, and it is known that fixation durations tend to be longer compared to silent reading, particularly in cases in which reading aloud is not a practiced skill (Rayner, 1998).

1.3. The present study

In the current study, non-university-level students and university-level students read a series of sentences, while their eye movements were monitored. Half of the sentences were accompanied by a related image, while the other half were not. In addition, we manipulated one word in the sentence, so that half of the sentences contained a low-frequency (target) word, and the other half contained a high-frequency (target) word. These manipulations were implemented in a blocked design, in which participants read 15 sentences in each of the four blocks in each condition. Participants' understanding of each sentence was assessed via a yes/no (true/false) comprehension question. Our hypotheses were formulated around on three main research questions:

RQ1. Do visual support and lexical simplification facilitate sentence comprehension and processing?

RQ2. Is there an interaction between visual support and lexical simplification?

RQ3. Are the effects of visual support or lexical simplification moderated by education level (i.e. is there an interaction between visual support or lexical simplification and education level) ?

We measured the impact of visual support and lexical simplification on comprehension accuracy (where we expected a positive impact of both inclusion of visual support and lower word frequency), but also on eye-tracking measures at the sentence and target levels (see the Method section for the specific variables). Reading times for a word before continuing to read the rest of the sentence (gaze durations) are indicative of the earlier processing and lexical access to each word. Total fixation durations for the sentences or for the target are considered to be an indicator of the processing the words in the sentence and integrating their meaning into an overall mental representation of the text. Regressions (returning to a word after it has been read for the first time) are an indicator of how easy or difficult it may be to integrate the word into the text.

For RQ1, we hypothesized that visual support would facilitate the processing of sentences, resulting in faster total reading times and fewer regressions across the text. Although some studies have found no significant differences with co-present visual support, such as images or symbols (Dye et al., 2007; Poncelas & Murphy, 2007), we expected differences in this study because the images were visual representations of the content of the sentences, and this should facilitate the construction of the mental representation of the sentence meaning. We also expected a possible impact on early word processing, with the presence of pictures reducing gaze duration. Regarding lexical simplification, we hypothesized that high-frequency words would also facilitate reading. We expected a clear impact on reducing gaze duration in the target word. This is a well-documented effect in skilled readers. Since lexical simplification affects only the target word, we did not expect reductions in early processing times (gaze duration) for the sentence as a whole. However, more frequent vocabulary, albeit in one word in the sentence, could facilitate the construction of sentence meaning, and thus reduce global sentence reading time and the proportion of regressions also across the sentence. These hypotheses were consistent with Rivero-Contreras et al. (2021), who found that both adaptations could benefit readers with different levels of print exposure and vocabularies (with medium and large effect sizes for visual support, η^2_p, between .11 and .43; and large effect sizes for lexical simplification, η^2_p between .17 and .58), in a study including dyslexic readers. It remained to be seen whether this effect was also observed in readers with lower education, although we expected it to be the case.

With respect to RQ2, we hypothesized that the facilitation effect of visual support for the target words would be greater with low-frequency words (i.e., with co-present visual support, the difference between high- and low-frequency words would be reduced). This effect was expected both for early (gaze duration) and later processing measures. This is because the image will enable comprehension of at least "gist"-type content and can do so quite quickly. Castelhano and Henderson (2008) explored how quickly information is extracted from a scene to activate gist through a brief presentation of a scene followed by the name of a target object. The target object was either consistent or inconsistent with the scene gist, but was never actually present in the scene. Scene gist activation was assessed as the degree to which participants responded "yes" to consistent versus inconsistent objects (response bias produced by scene gist). The results showed that the scene gist was activated after an exposure of only 42 ms and that the strength of the activation increased with longer presentation durations. Thus, in the present study, we expected that the effect of the inclusion of an unknown or otherwise difficult-to-process word would be reduced (if not eliminated) when texts were accompanied by pictures. In contrast, with no image present, we expected a standard word-frequency effect. In relation to RQ3, we hypothesized that the visual support and lexical simplification would benefit non-university adults more than the university adults, across all measures. Considering that vocabulary is a significant predictor of reading comprehension, particularly in adult struggling readers (Tighe & Schatschneider, 2016b) and that less proficient adult readers do not use phonological codes for word recognition (Chace et al., 2005) or do so less efficiently (Binder & Borecki, 2008), we expected that these two adaptations, individually, would each provide greater support for participants with a lower educational level and would make up for the vocabulary and other language limitations they may have due to less reading and educational experience. Thus, we expected more reduced reading times and regressions, as well as increased comprehension in the non-university group in response to pictures and higher word frequency. We did not have specific hypotheses with respect to a possible three-way interaction between the picture and word-frequency conditions and educational level.

2. Method

2.1. Participants

Sixty Spanish students aged between 16 and 58 years took part in this study. The sample was divided into two groups (see Table 1). The first group was composed of 30 students of adult-education schools in rural areas (N–UNI), who were studying for a *General Certificate of Secondary Education* (GCSE), and vocational schools, where only a few already held

Table 1
Demographic data and intelligence broken down by group.

	N–UNI (N = 30)		UNI (N = 30)		*p-value*
	M (SD)	Range	*M (SD)*	Range	
Age (years)	28.16 (11.08)	16.70–48.40	26.30 (7.65)	20.60–58.80	.45
Years of schooling	12.10 (2.22)	8–17	18.77 (2.84)	15–25	<.001
K-BIT (verbal IQ)	94.43 (9.57)	78–110	111.5 (7.87)	94–133	<.001
K-BIT (non-verbal IQ)	88.40 (10.32)	69–111	112.7 (6.76)	100–126	<.001
K-BIT (total IQ)	87.80 (8.29)	74–107	111.3 (6.85)	97–131	<.001
DELE (score)	4.93 (2.02)	1–9	9.63 (1.22)	7–12	<.001

N = number of participants; M = Mean; SD = Standard Deviation; IQ = intelligence quotient; N–UNI = students of adult schools; UNI = university students; K-BIT = Brief intelligence test of Kaufman; DELE = Spanish as a Foreign Language Diplomas.

the GCSE. The second group was comprised 30 university-level students (UNI), who were in the process of obtaining an undergraduate, Masters, or PhD degree. These students were recruited in a School of Psychology. Participants who were not native Spanish speakers were excluded. So were participants with intellectual disability, neurobiological/neurodevelopmental disorder, or other difficulty related to reading ability (e.g., dyslexia, speech disorders) and sensory disorders (e.g., hearing impairment). The two groups were matched in terms of gender (21 men and 9 women in both groups) and age.

The recruitment and selection process were as follows: first, the first author attended the classrooms of the educational centres. There, she reported on the study and asked for participation. Students who wanted to participate provided their contact details for later contact. Only one participant was not selected on suspicion of having an intellectual disability.

The Ethical Committee in Regional Biomedical Research of Andalusia approved the procedures for recruitment and collection of data. Informed written consent of the participants was obtained before the tests were carried out.

Sample size was restricted by availability and access to age-matched non-university participants. A sensitivity power analysis for 30 participants per group, with a .05 level of significance (alpha level) and power of 0.80 was computed using MorePower 6.0 (Campbell & Thompson, 2012). The minimal statistically detectable effect for a three-way mixed interaction was calculated at a partial eta squared of 0.13, a large effect size (for a medium effect size of $\eta_p^2 = .06$, power was reduced to 0.47).

2.2. Materials

2.2.1. Background measures

The following evaluation instruments that were used to characterize the sample:

The **Kaufman Brief Intelligence Test (K-BIT)** measures general intelligence of people aged 4–90 years (Kaufman et al., 2000). It consists of two subtests (vocabulary and matrices), providing verbal and nonverbal intelligence scores, and a combined score summarizing overall performance. Internal consistency reliabilities for ages 15–90 averages 0.92 for the overall K-BIT IQ Composite (from 0.83 to 0.96), 0.80 for the Vocabulary subtest – Verbal IQ (from 0.76 to 0.95) and 0.89 for the Matrices subtest – Non-verbal IQ (from 0.82 to 0.93). Test administration takes approximately 20–30 min.

Spanish as a Foreign Language Diplomas (Diplomas de Español como Lengua Extranjera - DELE) evaluates the level of Spanish. It includes four sections: reading comprehension, expression and written interaction, listening comprehension, and oral expression and interaction. For this study, Tasks 1 and 3 of the reading comprehension subtest were used. These are composed of two texts with six multiple-choice questions, producing a total score based on the number of correct answers. The internal consistency reliability of DELE was 0.75. Administering this test takes approximately 20–25 min (Instituto Cervantes, 2014).

2.2.2. Experimental manipulation

The **Sentence Processing Task** used a 2 × 2 × 2 (Picture: present vs. absent, Frequency: high vs. low, Group: non-university vs. university) mixed design, in which Picture and Frequency were within-subjects and Group was between-subject factors. Participants completed two practice trials and 60 experimental trials (15 per condition), which were presented in a random order (within each block) for each participant (see Fig. 1). The frequency of target words was established using the Spanish word database LEXESP (Sebastián, 2000). The frequencies-per-million mean of low-frequency (target) words in the corpus was 3.20 (SD = 3.60). For high-frequency words the frequencies-per-million mean was 33.55 (SD = 43.03). This difference was statistically significant $t(59) = -5.51$, $p < .001$, $d = -0.99$. The low- and high-frequency target words were similar in length ($n \pm 1$), and nouns, adjectives, and verbs were used equally. Sentences ranged in length from 15 to 22 words (M = 17.90). The only difference between sentence pairs was the frequency of the target word, since the manipulation only changed this word. These sentences had one or two subordinate clauses. The images, which were a visual representation of the sentences, were created by a graphic designer once the sentences were validated in the Sentence Norming Pre-Study (next section). All the main elements of the sentences were visually represented in the scene (i.e., contextualized), including in the target word. Words that expressed emotions were visually represented with the non-verbal/gestural language of the protagonist. The images, when present, were positioned below the sentence and centered on the screen (see Fig. 1). After each sentence, participants were presented with

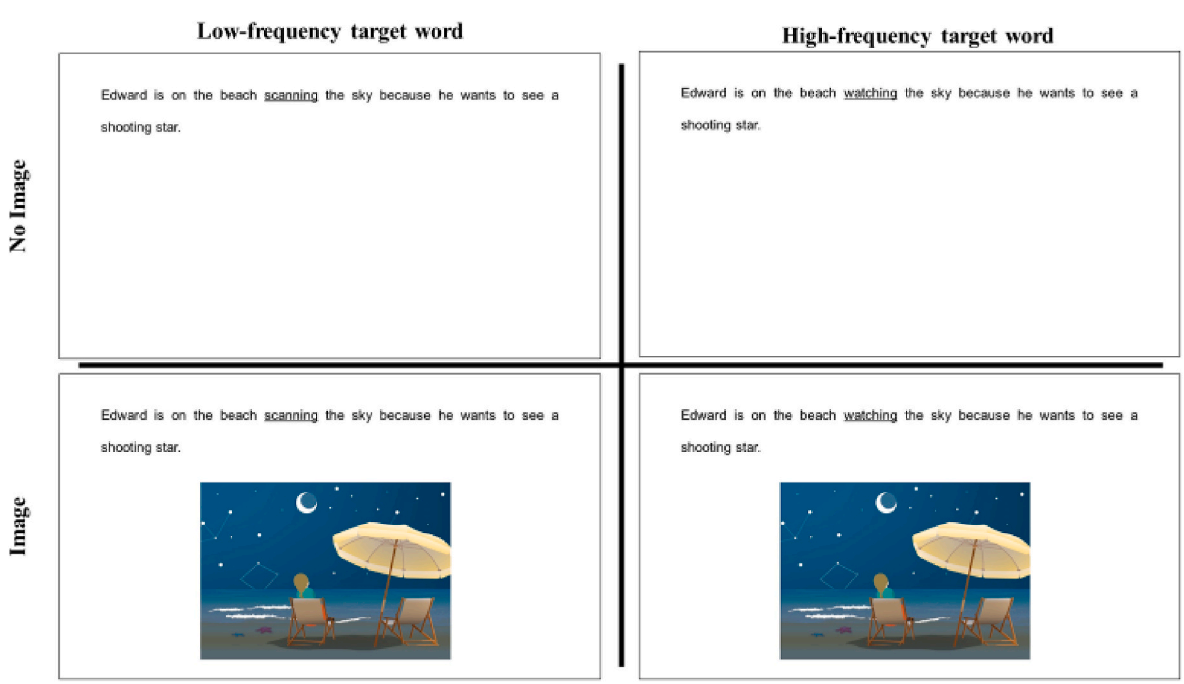

Fig. 1. Example item for each of the four within subjects conditions. The underlined word is the target word.

an auditory inferential statement, in which they needed to indicate whether the statement was "true" or "false" based on the preceding sentence. Internal consistency reliability for the task was 0.70 (0.68 for statements of sentences with a low-frequency word; 0.73 for statements of sentences with a high-frequency word; 0.71 for statements of sentences without an image; 0.66 for statements of sentences with an image). Following each block of 15 trials, participants were asked to indicate the degree of difficulty ("How easy has it been to understand these sentences?") on a scale from 1 (*very easy*) to 5 (*very difficult*). The sentence processing task took approximately 20–30 min to complete.

Sentence Norming Pre-Study. Prior to the study, the sentences were validated by 100 students from the University of Seville. Two versions of each sentence were presented: original (low frequency) and with lexical simplification (high frequency). The students sorted each pair in terms of difficulty ($1 = easy$, $2 = difficult$). Results showed that sentences with lexical simplification ($M = 1.49$, $SD = 0.13$) were rated significantly easier than the original sentences ($M = 1.91$, $SD = 0.09$). This difference was statistically significant $t(59) = 17.91$, $p < .001$, $d = -3.76$.

2.3. Procedure

The sentence processing task was carried out individually. Each participant was assigned to one of four lists so that s/he would only read each sentence once. First, participants read a set of instructions with the details of the procedure. Next, two example trials were completed. Following this, a nine-point calibration and validation procedure was carried out. The participant fixated nine points on the screen, and a maximum of 1.00 degree of deviation between point and record was required before they could proceed. Calibration and validation were repeated before each block, which allowed rest breaks during the experiment.

The sentence was presented on the top on the screen. If that trial contained an image, it was presented simultaneously below the sentence and centered. When the participant finished reading, s/he pressed the spacebar and a statement was presented auditorily, and participants then answered true or false to the statement (see Fig. 2). After completing 15 trials in a block, participants were asked to indicate how difficult they found the sentences in that block, and then the first trial of the next condition began. The standardized tests were administered before the sentence processing task, resulting in a testing session of approximately 70 min in total.

2.4. Apparatus

During the sentence processing task, an Eyelink 1000 plus was used (SR Research), with monocular recording (right eye) and sampling at 2000Hz. Participants used a chin rest to maintain head stability. Stimuli were presented on a computer screen of 21.5 inches (50 × 37.1 cm), positioned 55 cm from the participant. The sentences were presented in black font (Arial size 20) on a white background. The images were in colour. A keyboard, with marked response keys, was used to answer the comprehension questions.

2.5. Measurements

Comprehension Accuracy. It was the percentage of true/false statements answered correctly in each block of sentences.

Image Looking Time. It was the total time spent looking at the image area.

Target Word (Local) Measures. There were four target-word measures. *Total fixation duration* is the sum of the duration of all fixations on the word in milliseconds, including regressions back to the word. *Gaze duration* is the sum of all fixations on a word in milliseconds, excluding any fixations after the eye-gaze has left the word. These fixation duration measures are indicators of lexical access of word recognition (Inhoff & Rayner, 1986; Morton, 1969; Whaley, 1978) and are related to processing and integration. *Regressions* are indicative of problems with linguistic processing (Reichle et al., 2003). *Regression path duration* is the sum of all fixations in milliseconds from first entering a region until moving to the right of that region and is another indicator of processing difficulty (Hyönä et al., 2003).

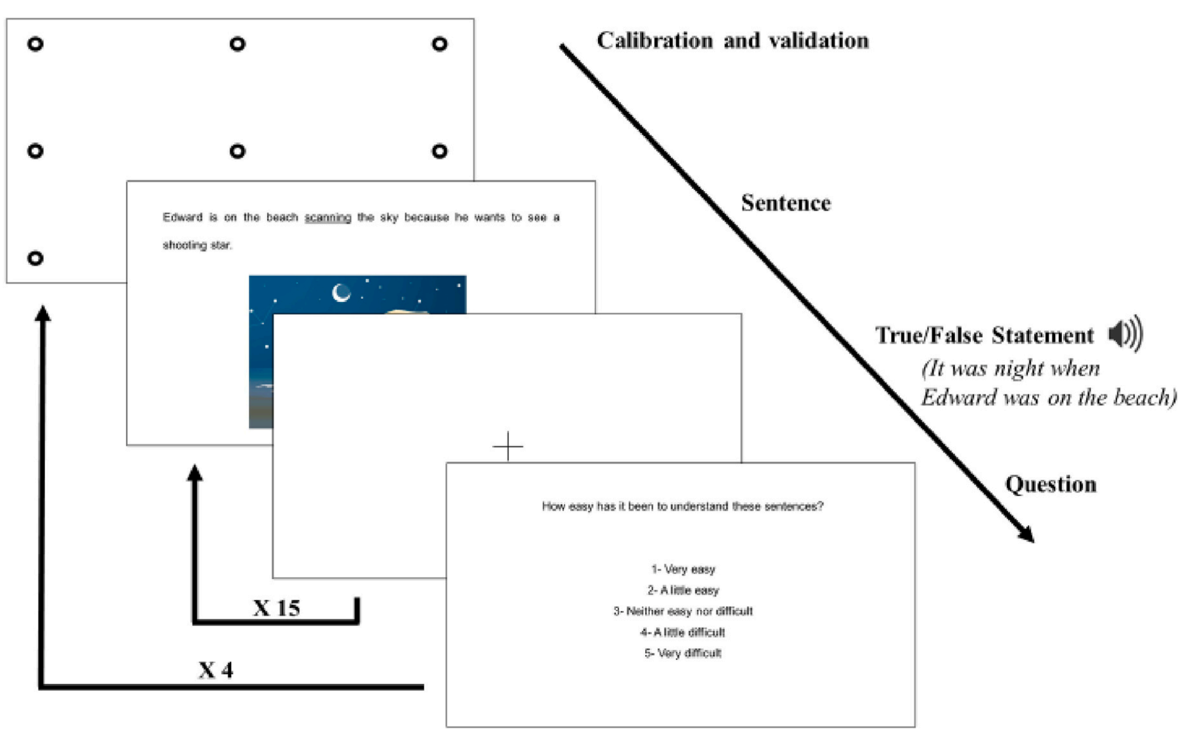

Fig. 2. Order of events within each block of the sentence processing task. There were 15 trials (X 15) within each block and each participant completed four blocks (X 4).

Sentence/Text (Global) Measures. There were three text-level measures (Text total fixation duration, Text gaze duration, and probability of Text regressions). Text total fixation duration and Text gaze duration were the sum of all word reading times and word gaze durations on the sentence in seconds, respectively. These times did not include time dedicated to looking at the pictures. To calculate the probability of Text regressions in terms of 0–1, we summed all probabilities of regressions out of each area of interest on the sentence. Then, we divided the total probability of regressions out of each word by the number of words in each sentence. Finally, we calculated the mean of the probability of target regressions in each block.

Perceptions of Text Difficulty. It was participants' perceptions of text ease/difficulty on a scale of 1–5 (1 = very easy; 5 = very difficult). It was included as a secondary measure in this study since perceived text difficulty by target readers is recommended to test text adaptions in most guidelines.

2.6. Data screening and analysis

Correlations between the individual characteristics (Table 1) and the dependent variables (eye-movement measures) for both groups are reported in supplementary material (Tables s1, s2 and s3).

Fixations of less than 80 ms and longer than 1200 ms were excluded from the dataset. Data from each sentence were reviewed, and items that were not recorded or that contained excessive blinks (more than 50% of trial not recorded) were excluded from the analyses, resulting in data loss of 4.17%. One university student was excluded for excessive blinks. Data for each eye-movement measure greater than 3 SDs from the mean were defined as outliers. Outliers represented 2.03% of the data and were replaced with the mean of that variable (McCartney et al., 2006).

Nonparametric analyses across subjects and items were performed on comprehension accuracy. Main effects of Picture and Group by subjects were assessed using Wilcoxon Signed-Rank Test (Z_1) and Mann-Whitney *U* Test (U_1), respectively. Main effects of Picture and Group by items were assessed using Wilcoxon Signed-Rank Test (Z_2). The Frequency × Picture interaction was probed using Wilcoxon Signed-Rank Test across subjects (Z_1) and items (Z_2). Bonferroni correction was applied to allow for multiple comparisons (p's < 0.0125 were considered significant).

Image looking times were assessed using an independent-samples *t*-test to compare the means of Group by subjects (t_1) and paired-samples *t*-tests to compare the means of Frequency by subjects (t_1) and items (t_2), and Group by items (t_2).

We analysed the global and local measures data using mixed ANOVAs for each of the measures (Picture × Frequency × Group) across subjects (F_1) and items (F_2). Target regressions out and Regression path duration were log-transformed to meet the normality assumption. Independent-samples *t*-tests were used to compare the means in Frequency × Group and Picture × Group interactions by subjects (t_1) and paired-samples *t*-tests were used to compare the means in three interactions by items (t_2). Moreover, in Target total fixation duration and Target gaze duration we used independent-samples *t*-tests to compare the differences of means of frequencies in Frequency × Group interactions by subjects (t_1) and paired-samples *t*-tests to compare the differences of means of frequencies in this interaction by items (t_2). FDR corrected significance levels were also used to interpret the level of significance of effects sizes across the different ANOVAs used for the different eye-tracking measures (particularly for the main effects of the Picture condition; supplementary Table s7). Bonferroni corrections were applied to allow for multiple comparisons in post-hoc follow up of significant interactions (p's < 0.0125 were considered significant).

The participants' perceptions of text ease/difficulty were analysed using a 2 × 2 × 2 (Picture × Frequency × Group) mixed ANOVA.

False discovery rate (FDR) corrections were applied to check correlations between Image Looking Time and of eye movement measures of Picture condition, Comprehension Accuracy and Perceptions of Text Difficulty.

We report η_p^2, *r,* and *Cohen's d* for effect sizes (Field, 2009).

3. Results

3.1. Comprehension accuracy

There were significant main effects of Picture and Group in comprehension accuracy (see Table 2). Sentences with an image had higher accuracy than sentences without an image, and the university group had greater precision compared to the non-university group. There was not a significant main effect of Frequency (see Table 2). In addition, paired comparisons for the Frequency × Picture interaction revealed that the comparison of low-vs. high-frequency words with an image and no image were not significant in either case. But the comparison of image present vs. image absent was significant for low-frequency words, although not for high-frequency words (see Table 2).

3.2. Image looking time

There were no significant main effects of Frequency or Group on image looking time (although t_2 analyses showed a significant a main effect of Group) (see Table 2).

Only two negative correlations were found in Total Gaze Duration (Picture) and Perceptions of Text Difficulty with Image Looking Time (see supplementary Table s5); however, the p-values were not significant when the FDR was applied (see supplementary Table s6).

The lack of effects or correlations implicating image looking time appear to indicate that the time spent on viewing the picture is not directly related to comprehension or reading processing. No further considerations or analyses were therefore made involving image processing.

3.3. Global measures

For the three Text measures (Text total fixation duration, Text gaze duration and Text regressions) we found a main effect of Picture, but only for Text total fixation duration and Text gaze duration we observed a main effect of Group. Sentences with an image showed shorter total fixation durations, gaze durations, and a less probability of regressions than sentences without an image; and university adults had shorter total fixation durations and gaze durations than non-university adults. No other main effects or interactions were significant (see Table 3 for the means and Table 4 for the full ANOVA).

3.4. Local measures

For the four Target word measures, we found mains effect of Picture and Frequency, but only for measures Target total fixation duration, Target gaze duration and Regressions path duration did we observe a main effect of Group (see Table 5 for the means and Table 6 for the full ANOVA). Sentences with an image showed shorter total fixation durations, gaze durations, regression path durations, and a less probability of regressions out than sentences without an image; the high-frequency words had shorter total fixation durations, gaze durations, regression path durations, and a less probability of regressions out than low-frequency words; and the university group had shorter total fixation durations, gaze durations, regression path durations compared to the non-university group.

Also, we found a Frequency × Picture interaction in Target gaze duration and Regression path duration (see Table 6).

Paired comparisons for the Frequency × Picture interaction in Target gaze duration revealed that the comparison of low-vs. high-frequency words was significant both when the picture was absent and present (see Table 7). The paired comparison of low-frequency words with image present vs. image absent was also significant; however, the paired

Table 2
Mann-whitney *U* test and wilcoxon signed-rank test of comprehension accuracy.

	M (SD)	M (SD)	Z_1/U_1	p_1	r_1	Z_2/U_2	p_2	r_2
Comprehension Accuracy								
·Picture (Absent vs. Present)	83.78 (12.08)	86.56 (9.61)	−2.28	.023	0.29	−2.28	0.23	0.29
Frequency (Low vs. High)	85.33 (10.70)	85.00 (10.73)	−0.27	.788	0.02	−0.27	.788	0.02
·Group (N–UNI vs. UNI)	80.78 (10.67)	89.56 (6.59)	167	<.001	0.54	−3.42	.001	0.62
Frequency (Low vs. High) x Picture (Absent)			−1.07	.286	0.14	−1.72	.085	0.22
Frequency (Low vs. High) x Picture (Present)			−1.45	.147	0.19	−0.74	.460	0.10
·Frequency (Low) x Picture (Absent vs. Present)			−3.01	.003	0.39	−2.86	.004	0.37
Frequency (High) x Picture (Absent vs. Present)			−0.34	.737	0.10	−0.09	.928	0.01
	M (SD)	***M (SD)***	***t_1 (58)***	***p_1***	***d_1***	***t_2 (59)***	***p_2***	***d_2***
Image Looking Time								
Frequency (Low vs. High)	432.36 (281.08)	433.95 (272.53)	−0.57	.955	−0.01	−0.80	.425	−0.13
Group (N–UNI vs. UNI)	475.43 (253.59)	392.30 (253.53)	1.23	.816	0.32	−5.34	<.001	−0.68

Note. · significant effects.

Table 3
Means (and standard deviations) of global eye-tracking measures.

	Non-University group							
	No Picture – Low		Picture - Low		No Picture - High		Picture - High	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range	*M (SD)*	Range
Measure								
Text total fixation duration (sec)	88.57 (27.63)	36.40–151.74	86.08 (28.32)	51.26–177.98	89.70 (33.27)	50.55–168.27	84.53 (30.87)	27.19–167.77
Text gaze duration (sec)	62.73 (14.52)	32.40–91.56	61.88 (15.58)	41.98–109.77	61.78 (15.20)	38.82–100.44	59.76 (15.88)	23.44–94.11
Text regressions (p)	0.09 (0.04)	0.01–0.20	0.08 (0.05)	0.00–0.22	0.10 (0.05)	0.02–0.19	0.08 (0.05)	0.01–0.19
	University group							
	No Picture – Low		**Picture - Low**		**No Picture - High**		**Picture - High**	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range	*M (SD)*	Range
Measure								
Text total fixation duration (sec)	66.03 (16.61)	39.42–105.82	65.43 (19.61)	33.40–125.22	68.20 (18.77)	32.40–121.74	61.99 (19.86)	28.97–122.40
Text gaze duration (sec)	46.75 (8.58)	28.59–70.86	45.01 (8.83)	25.13–66.50	45.51 (8.37)	26.16–67.71	43.07 (9.33)	22.20–64.46
Text regressions (p)	0.11 (0.05)	0.04–0.22	0.10 (0.04)	0.05–0.22	0.10 (0.05)	0.03–0.23	0.10 (0.05)	0.00–0.20

Table 4
ANOVAs analysis of Global Reading Time Measures.

	M (SD)	*M (SD)*	*F_1* (1,57)	*p_1*	η^2_{p1}	*F_2* (1,59)	*p_2*	η^2_{p2}
			Text total fixation duration					
Picture (Absent vs. Present)	78.31 (26.09)	74.69 (25.85)	6.88	.011	.11	16.65	<.001	.22
Frequency (Low vs. High)	76.71 (25.04)	76.29 (27.03)	0.08	.774	.00	1.38	.244	.02
Group (N–UNI vs. UNI)	87.22 (27.47)	65.41 (17.55)	13.10	.001	.19	1017.37	<.001	.95
Picture x Group			0.02	.880	.00	0.73	.395	.01
Frequency x Group			0.02	.886	.00	0.60	.441	.01
Frequency x Picture			1.65	.204	.03	0.84	.363	.01
Frequency x Picture x Group			0.21	.651	.00	0.20	.660	.00
			Text gaze duration					
·Picture (Absent vs. Present)	54.33 (14.01)	52.50 (15.03)	6.16	.016	.10	7.57	.008	.11
Frequency (Low vs. High)	54.23 (14.03)	52.67 (14.84)	7.17	.010	.11	2.27	.137	.04
Group (N–UNI vs. UNI)	61.54 (14.25)	45.08 (8.31)	29.09	<.001	.34	1263.97	<.001	.96
Picture x Group			0.22	.644	.00	0.00	.956	.00
Frequency x Group			0.00	.965	.00	0.17	.680	.00
Frequency x Picture			0.44	.508	.01	0.09	.763	.00
Frequency x Picture x Group			0.03	.867	.00	0.00	.970	.00
			Text regressions					
Picture (Absent vs. Present)	0.10 (0.04)	0.09 (0.04)	8.07	.006	.12	14.58	<.001	.20
Frequency (Low vs. High)	0.10 (0.04)	0.10 (0.04)	0.01	.910	.00	0.03	.874	.00
Group (N–UNI vs. UNI)	0.09 (0.05)	0.10 (0.05)	1.40	.241	.02	0.30	.584	.01
Picture x Group			0.96	.332	.02	1.77	.189	.03
Frequency x *Group*			2.18	.146	.04	1.22	.274	.02
Frequency x Picture			0.57	.453	.01	0.96	.331	.02
Frequency x Picture x Group			0.18	.677	.00	0.05	.822	.00

Note. · significant effects.

Table 5
Means (and standard deviations) of local eye-tracking measures.

	Non-University group							
	No Picture – Low		Picture - Low		No Picture - High		Picture - High	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range	*M (SD)*	Range
Measure								
Image Looking Time (ms)	–	–	488 (286)	62–973	–	–	463 (269)	12–1066
Target total fixation duration (ms)	993 (380)	482–2082	861 (364)	314–1742	609 (268)	255–1266	589 (213)	258–1163
Target gaze duration (ms)	611 (210)	275–1204	518 (209)	234–1111	372 (106)	221–617	388 (106)	229–623
Target regressions out (p)	0.16 (0.09)	0.00–0.36	0.14 (0.12)	0.00–0.47	0.17 (0.11)	0.00–0.40	0.14 (0.12)	0.00–0.44
Regression path duration (ms)	1030 (406)	450–2177	802 (327)	258–1554	577 (185)	221–960	599 (245)	266–1258
	University group							
	No Picture - Low		**Picture - Low**		**No Picture - High**		**Picture - High**	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range	*M (SD)*	Range
Measure								
Image Looking Time (ms)	–	–	379 (270)	0–1057	–	–	406 (277)	47–1089
Target Total fixation duration (ms)	572 (162)	270–912	526 (184)	251–998	424 (126)	197–691	384 (137)	174–782
Target gaze duration (ms)	347 (91)	186–584	314 (127)	212–798	279 (57)	175–447	267 (60)	164–446
Target regressions out (p)	0.20 (0.13)	0.00–0.46	0.17 (0.10)	0.00–0.33	0.16 (0.13)	0.00–0.43	0.10 (0.08)	0.00–0.29
Regression path duration (ms)	628 (261)	278–1337	499 (231)	261–1297	405 (114)	236–645	390 (164)	184–1101

Table 6
ANOVAs analysis of Local Reading Time Measures.

	M (SD)	*M (SD)*	F_1 (1,57)	p_1	η^2_{p1}	F_2 (1,59)	p_2	η^2_{p2}
			Target total fixation duration					
Picture (Absent vs. Present)	651 (280)	592 (252)	12.21	.001	.18	10.40	.002	.15
Frequency (Low vs. High)	489 (206)	503 (208)	122.17	<.001	.68	171.32	<.001	.74
Group (N–UNI vs. UNI)	763 (277)	477 (130)	25.60	<.001	.31	417.11	<.001	.88
Picture x Group			0.96	.331	.02	1.72	.195	.03
Frequency x Group			18.33	<.001	.24	47.11	<.001	.44
Frequency x Picture			3.71	.059	.06	1.86	.178	.03
Frequency x Picture x Group			3.09	.084	.05	1.27	.265	.02
			Target gaze duration					
Picture (Absent vs. Present)	403 (142)	373 (144)	5.99	.017	.10	5.79	.019	.09
Frequency (Low vs. High)	503 (208)	448 (184)	105.52	<.001	.65	236.48	<.001	.80
Group (N–UNI vs. UNI)	472 (134)	302 (68)	37.49	<.001	.40	377.63	<.001	.87
Picture x Group			0.38	.540	.01	1.01	.319	.02
Frequency x Group			29.07	<.001	.34	49.27	<.001	.23
Frequency x Picture			7.08	.010	.11	17.67	<.001	.23
Frequency x Picture x Group			3.38	.071	.06	5.76	.020	.09
			Target regressions out[a]					
Picture (Absent vs. Present)	0.17 (0.10)	0.14 (0.08)	7.56	.008	.12	5.66	.021	.09
Frequency (Low vs. High)	0.17 (0.09)	0.14 (0.09)	5.76	.020	.10	7.63	.008	.12
Group (N–UNI vs. UNI)	0.15 (0.11)	0.16 (0.11)	0.10	.752	.00	0.02	.895	.00
Picture x Group			0.18	.673	.00	0.04	.841	.00
·Frequency x Group			6.10	.017	.10	7.27	.009	.11
Frequency x Picture			0.46	.499	.01	0.02	.887	.00
Frequency x Picture x Group			0.10	.752	.00	0.04	.849	.00
			Regression path duration[a]					
Picture (Absent vs. Present)	660 (256)	574 (241)	13.84	<.001	.20	12.45	.001	.17
Frequency (Low vs. High)	740 (309)	494 (173)	179.13	<.001	.76	207.44	<.001	.78
Group (N–UNI vs. UNI)	752 (233)	481 (127)	34.63	<.001	.38	391.22	<.001	.87
Picture x Group			0.12	.728	.00	0.58	.450	.01
Frequency x Group			3.89	.053	.06	1.76	.190	.03
·Frequency x Picture			8.03	.006	.12	13.38	.001	.19
Frequency x Picture x Group			0.50	.484	.01	0.24	.624	.00

Note. · significant effects.
[a] The data set of items (F_2) did not have a normal distribution. Thus, non-parametric analyses were performed. The same main effects and interaction were obtained.

comparison of high-frequency words with image present vs. image absent was not significant (see Fig. 3) (see Table 7).

In Regression path duration, the paired comparisons of low-vs. high-frequency words with an image and no image were significant (see Table 7). The comparison of picture present vs. absent was significant for low-frequency, but not for the high-frequency words (see Fig. 4) (see Table 7).

In addition, we observed a Frequency × Group interaction in three Target word measures (see Table 6).

In Target total fixation duration, although all paired comparisons were significant ($p < .001$), the comparison of the differences in the frequencies of target words between groups was significant (non-university had a larger difference than university), $t_1(35.77) = 4.38$, $p < .001$, $d = 1.46$; $t_2(59) = 7.34$, $p < .001$, $d = 1.24$ ($M_{N\text{-}UNI} = 328.63$, $SD_{N\text{-}UNI} = 218.76$; $M_{UNI} = 143.69$, $SD_{UNI} = 74.11$), which explained the interaction (see Fig. 3). But the adaptation did benefit both groups: frequency-based comparisons revealed that low-frequency words had longer total fixation durations than high-frequency words in both

Table 7
Interactions analysis of Local Reading Time Measures.

	M (SD)	*M (SD)*	gl_1	t_1	p_1	d_1	gl_2	t_2	p_2	d_2
Target gaze duration										
Frequency (Low vs. High) x Picture (Absent)	481.03 (209.42)	326.39 (96.72)	58	5.08	<.001	0.66	59	7.78	<.001	1.00
Frequency (Low vs. High) x Picture (Present)	418.03 (200.90)	328.60 (105.61)	58	7.67	<.001	1.00	59	11.65	<.001	1.50
Frequency (Low) x Picture (Absent vs. Present)	481.03 (209.42)	418.03 (200.90)	58	2.69	.009	0.35	59	3.66	.001	0.47
Frequency (High) x Picture (Absent vs. Present)	326.39 (96.72)	328.60 (105.61)	58	−0.29	.771	0.04	59	−0.48	.636	0.06
Target regressions out										
Frequency (Low vs. High) x Group (UNI)	0.19 (0.09)	0.13 (0.08)	28	3.55	.001	0.66	59	3.47	.001	0.45
Frequency (Low vs. High) x Group (N–UNI)	0.15 (0.09)	0.15 (0.09)	29	−0.07	.952	0.01	59	−0.23	.823	0.03
Frequency (Low) x Picture (UNI vs. N–UNI)	0.19 (0.09)	0.15 (0.09)	57	−1.42	.162	0.37	59	−1.65	.105	0.21
Frequency (High) x Picture (UNI vs. N–UNI)	0.13 (0.08)	0.15 (0.09)	57	1.01	.319	0.26	59	1.85	.070	0.24
Regression path duration										
Frequency (Low vs. High) x Picture (Absent)	832.48 (395.49)	492.64 (176.12)	58	5.09	<.001	1.30	59	11.50	<.001	1.48
Frequency (Low vs. High) x Picture (Present)	653.15 (320.60)	496.35 (232.50)	58	5.09	<.001	0.66	59	7.11	<.001	0.92
Frequency (Low) x Picture (Absent vs. Present)	832.48 (395.49)	653.15 (320.60)	58	4.51	<.001	0.59	59	4.44	<.001	0.57
Frequency (High) x Picture (Absent vs. Present)	492.64 (176.12)	496.35 (232.50)	58	0.36	.717	0.05	59	0.98	.334	0.13

Note. · significant effects.

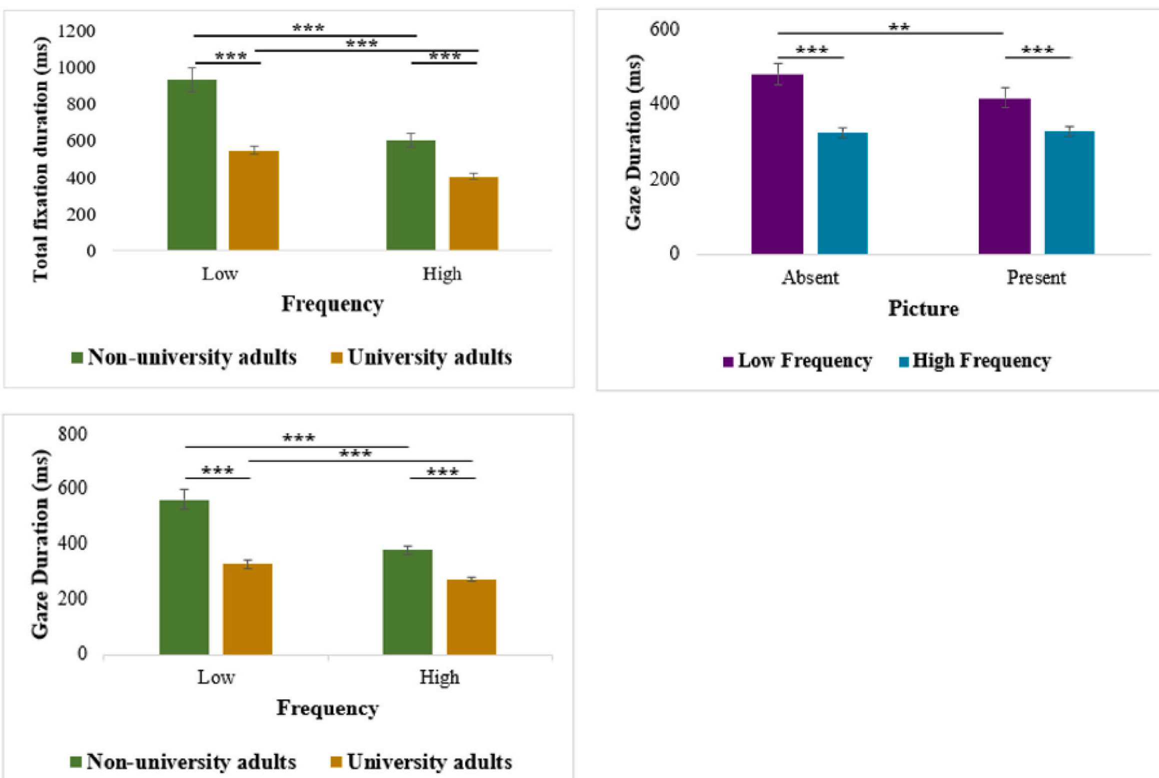

Fig. 3. Upper left panel shows the mean Target total fixation durations for the Frequency × Group interaction. Upper right panel shows the mean Target gaze duration for the Frequency × Picture interaction. Lower left panel show mean Target gaze duration for the Frequency × Group interaction. Error bars indicate standard error of the mean. *** $p < .001$, ** $p < .01$.

participant groups [non-university: $t_1(29) = 8.23$, $p < .001$, $d = 1.50$; $t_2(59) = 12.06$, $p < .001$, $d = 1.56$; university: $t_1(28) = 10.44$, $p < .001$, $d = 1.94$; $t_2(59) = 7.53$, $p < .001$, $d = 0.97$] (see Fig. 3).

Paired comparisons for the Frequency × Group interaction in Target gaze duration were all significant. The low-frequency words resulted in longer Target gaze duration than high-frequency words for both groups [non-university: $t_1(29) = 8.68$, $p < .001$, $d = 1.58$; $t_2(59) = 16.19$, $p < .001$, $d = 2.09$; university: $t_1(28) = 5.74$, $p < .001$, $d = 1.07$; $t_2(59) = 11.13$, $p < .001$, $d = 1.44$]. But, once again, the effect was larger in the non-university group and the comparison of the differences in the frequencies of target words between groups was significant $t_1(40.78) = 5.47$, $p < .001$, $d = 1.71$; $t_2(59) = 8.08$, $p < .001$, $d = 1.62$ ($M_{N\text{-}UNI} = 115.74$, $SD_{N\text{-}UNI} = 21.13$; $M_{UNI} = 52.59$, $SD_{UNI} = 9.77$), as expected from the interaction (see Fig. 3).

In Target regressions out, paired comparison indicated no significant differences between the two groups for the sentences with low-frequency, or the sentences with high-frequency words. The non-university students revealed no difference based on word frequency, but the university students showed significantly greater probability of regressions for the sentences with low-frequency words (see Fig. 4) (see Table 7).

No other main effects or interactions were significant (see Table 6 for the full ANOVA).

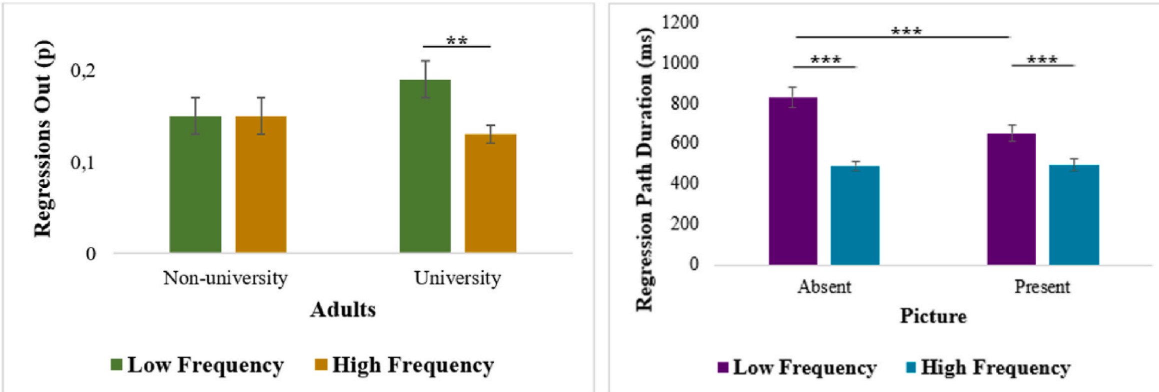

Fig. 4. Left panel shows the mean Target regressions for the Frequency × Group interaction. Right panel shows the mean regression path duration for the Frequency × Picture interaction. Error bars indicate standard error of the mean. *** $p < .001$, ** $p < .01$.

3.5. Perceptions of text difficulty

Participants' perceptions of text difficulty, assessed by the final question in each block, showed that participants evaluated sentences as relatively easy to understand (overall mean: $M = 2.43$, $SD = 0.79$). There were no main effects or interactions of Picture, Frequency, or Group (all p's > 0.05; see supplementary Table s4 for the full ANOVA).

4. Discussion

This study evaluated the efficacy of visual support and word frequency (i.e., *Easy-to-Read* suggestions) on sentence processing in adults with differing education levels.

The first research question, which focused on visual support and lexical simplification, led us to hypothesize that both would have a positive impact on processing and comprehension accuracy. To a large extent, this hypothesis was borne out in our data. In the following discussion, we focus primarily on the impact of visual support, given that word frequency has been amply demonstrated in many psycholinguistic studies (e.g., Ashby et al., 2005; Hyönä & Olson, 1995). The presence of a picture led to more accurate comprehension (a 6% difference) (see also Jones et al., 2007). With respect to facilitation of processing (eye tracking measures) at the text- and word-level, we found reduced Text gaze duration, Text total fixation duration, and probability of regression in sentences accompanied by an image. That is, picture produced a main effect in which presence of a picture reduced processing effort, and similar effects were observed at the word level. These results indicate that visual support assists in the processing of words at both the overall sentence level and at the level of individual words contained within the sentence. It also appears to facilitate both early and late processing. Since our visual support provided a broad depiction of the meaning of the sentence, we expected that participants would be able to extract situational-event information (i.e., propositional content), which would allow readers to predict either specific upcoming words or to activate certain regions in the semantic network. Thus, the eye-tracking data suggests that visual support assists in the creation/activation of mental representations, which can be combined (and integrated) together with the linguistic context, and thus, facilitates processing (Mayer, 2009; Sadoski & Paivio, 2004, 2012).

We acknowledge that the effect of visual support seems obvious and straightforward, as multiple streams of information or multiple avenues to meaning should always be beneficial. However, as noted in the Introduction, the literature on visual support is extremely mixed: (1) studies use different types of visual support such as images, graphics, and symbols; (2) studies with the same type of visual support produce different results; and (3) different effects of visual support are found in the population with and without disabilities. For instance, different results have been observed with supporting symbols (Jones et al., 2007; Poncelas & Murphy, 2007). In addition, in the meta-analysis by Guo et al. (2020), studies with adults with a medium-large effect size were identified with different types of support in black and white (picture and pictorial diagram) and in color (flow diagram) (i.e., Jian & Wu, 2015; Liu et al., 2009; Mayer et al., 1996; Mayer & Gallini, 1990; Waddill et al., 1988).

With the current study we show that visual support moderately helps in adult reading; however, we still need more research to know which type of visual support is the most useful to support reading, and whether the goal of the reading task is involved in the selection of the type of visual support. In any case, it seems that general contextualizing pictures which include all elements of a sentence can be moderately supportive of its comprehension.

The second research question focused on whether visual support and lexical simplification would interact to affect sentence processing. Obviously, the low-frequency/no picture condition is the most difficult, and the high-frequency/picture condition is the easiest. Interactions are essentially dependent upon the results of the other two middle-difficulty conditions (i.e., low-frequency/picture and high-frequency/no picture). There were no interactions in how difficult participants judged the sentences to be. But we did observe interactions between Frequency and Picture in accuracy and in the eye-movement measures. For low-frequency words, the presence of the picture increased comprehension accuracy. For both Target gaze duration and Target regression path durations, results showed that the presence/absence of the picture reduced processing times when the sentence contained a low-frequency word. In contrast, processing times with the high-frequency words were lower overall and not affected as much by the presence/absence of an image. Therefore, the presence of an image reduced the slower lexical access associated with recognition of low-frequency words. These results corroborate previous evidence of visual support facilitating lexical access, and the understanding of unfamiliar words (Huettig et al., 2011; Montag et al., 2015; Qu et al., 2016), but again, this result has not been consistent in the literature. Also consistent with the literature is our finding that the effect sizes for the use of pictures is small to moderate. Potentially, pictures may have a limited effect at this educational level. But the results emerging from our study and the literature in general could be less clearcut than for vocabulary because the relationship between the picture and the text is in general less straightforward. Participants can interpret and use visuals in many ways, and in the absence of specific instructions to use the picture (although they did all look at the images), they may be making an inefficient use of it (Jian, 2018). Across studies, pictures themselves bear very different relations with text and are of diverse perceptual quality and content. Although careful consideration was given to including target word content in the pictures, other visual elements may have been less well controlled for.

The third research question focused on the effects of education level on sentence processing facilitation of picture and word frequency. Related to the effect of educational level on sentence reading as a whole, and as expected, we found that non-university students comprehended sentences less accurately than university students. We found a 10% difference in comprehension accuracy, which represents a large effect size ($r = 0.54$). This is similar to data reported by PIACC (Ministerio de Educación, Cultura y Deporte, 2013), in which lower reading comprehension is observed in individuals with low-to-medium education level. Educational level was up to a point confounded with IQ and language measures. There is most likely a mutual influence between these variables: greater exposure to formal education will impact positively on cognitive and linguistic development, but also cognitive and linguistic abilities will impact academic achievement (Ritchie & Tucker-Drob, 2018). In general, the non-university students spent longer reading the sentences, as shown by several significant main effects; and there were no differences between the groups in image processing times, as in the study by Hannus and Hyönä (1999). However, we cannot claim that there are different patterns of processing between the groups in relation to the presence of the image, as proposed in previous studies (Hannus & Hyönä, 1999; Jian, 2017; Mason et al., 2013b). To understand this issue in depth, we would need to analyse the transitions between text and image.

But our research question was focused on the interaction of our text adaptations with educational level (and not on these main effects). We had hypothesized that visual support and lexical simplification would differentially influence sentence processing in the two groups. The primary significant interactions involved word frequency. Across all the measures, we never observed an interaction of Picture and Group. The Frequency by Group interactions showed that lexical simplification (i.e., high-frequency words) resulted in Target word total fixation durations and gaze durations for the non-university students that were much more like the total fixation durations in university students. Thus, our data clearly shows that lexical simplification produces a strong positive effect on processing, specifically in individuals with a low level of education.

The interactions with word frequency are similar to Barnes et al. (2017), who explored the reading skills of a group of adults enrolled in a basic education program. The lower education adults exhibited longer fixations (like what we found) and a greater number of regressions (see also Ashby et al., 2005). In our study, we found differences in regression path duration but not in regressions out between groups (Barnes et al. did not examine regression path duration.) With respect to regressions, it is possible that the difference between low- and high-frequency words was such that it caused integration difficulty (requiring re-reading) for both groups of participants. Further work is needed to confirm this possibility.

However, despite there being no differences in the probability of regressions out, we found a Frequency by Group interaction effect in which university adults decreased probabilities of regressions out with lexical simplification while non-university adults did not. Comprehension monitoring refers to the identification and repair of misunderstandings when reading a text. Previous studies have demonstrated that elementary school students with greater literacy, academic language, and vocabulary implement comprehension repair strategies when inconsistent words are found in the text, resulting in more time rereading (Connor et al., 2015; Zargar et al., 2020). In our study, low-frequency words could have been perceived as inconsistencies in sentences, leading to the observed reduction in regressions in university adults when they were substituted for high-frequency words. In contrast, non-university adults, who also had a lower level of vocabulary, did not present these monitoring strategies (Tighe & Schatschneider, 2016b), and evidenced no change in probability of regressions out with lexical simplification.

4.1. Implications and applied value

The most obvious implication of our study for individuals with low education/poor comprehension is the usefulness of lexical simplification. Specifically, substituting high-frequency words for low-frequency words helped non-university adults reduce the processing time associated with lexical access and integration, enabling them to read more similarly to university adults. Therefore, our data supports the idea of using simple language, as suggested in the Easy-to-Read recommendations (Asociación Española de Normalización, 2018; Freyhoff et al., 1998; Tronbacke, 1997). Regarding visual support, we did not find differences between groups based on the presence or absence of an image. However, we did observe main effects of Picture in terms of both comprehension accuracy and eye tracking measures. In short, an image can facilitate processing and comprehension, and benefit everyone regardless of their education level. Beyond that, the effect of picture tended to be larger when the sentence contained a low-frequency word, in increasing comprehension accuracy and reducing Target gaze duration and Regression path duration more (but not so in the case of the rest of eye-tracking variables). Since the effect was not found across all measures, this result requires replication and should be treated with caution. But it points to the possibility that pictures could more useful in the context of low-frequency words and that, if so, designers should carefully consider the vocabulary specifically used in a text when drawing supporting illustrations.

Also, in the present study, our participants were unaware of the impact of visual support and lexical simplification on reading comprehension and sentence processing. Readers' perceptions of text difficulty are recommended and used as indicators of the impact of text simplification (Asociación Española de Normalización, 2018; Cerga-Pashoja et al., 2019; Karreman et al., 2007). Our data suggest that objective indicators of psycholinguistic processing, such as eye movement measures, should additionally be included to assess the degree of comprehension of easy-to-read written materials.

4.2. Limitations

The results of this study have yielded interesting and important findings. However, several limitations must be kept in mind. The first is that we used cartoon-type images as visual support, and so, at this time, we are limited in how much our results map onto the other studies in the literature that used photographs or symbols as visual support.

Second, the sentences that we tested were relatively simple overall, and resulted in generally high comprehension accuracy. There were some ceiling effects in the dataset (three participants), particularly within the university student sample.

Third, due to restrictions in access to the non-university participants, the sample size was relatively underpowered to reliably detect small to medium effect sizes. This limitation did probably not affect the main effects of interest (vocabulary, visuals, and educational level). Effects of interest in previous studies with this paradigm, specifically Rivero-Contreras et al. (2021), have found medium to large effects for the inclusion of pictures with text (with a minimum of .11, for which power in our study was 0.75) and large for lexical simplification. Studies comparing eye-tracking measures in groups with different educational levels found differences with large effect sizes (Barnes et al., 2017). In our case, it is thus unlikely that any of these main effects went undetected. But this might not be the case for the interactions, especially those including the presence of images. Conclusions regarding the lack of differential effects of these text adaptations should thus be treated with caution.

Finally, we feel that the timing of visual support is an important variable. In the current study, the image was presented simultaneously with the sentence, and so we do not know whether visual support would have made a larger impact on processing if it was presented prior to the sentence. This is an important issue, which should be addressed in future studies.

5. Conclusions

In summary, visual support facilitates sentence processing. This type of support promotes both lexical access and integration, and results in fewer regressions. Lexical simplification also helps processing and comprehension of sentences, and lexical simplification produces a much larger effect on individuals with a low level of education. Estimates suggest that one-fifth of the population of Spain meet the criteria for "low" education level, and thus, a sizable portion of people may benefit from Easy-to-Read suggestions in certain contexts. Therefore, the main take-home message of this study is that lexical simplification seems to benefit processing of written information for individuals with a low level of education, and the second take-home message is that visual support produces a more general effect.

Funding

This research was supported by the V Plan Propio de Investigación y Transferencia of the University of Seville [grant number PPI2016-IV.5]. The work of M. R. C. is supported by a research grant from the University of Seville [grant number PP2018-10696].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2022.101731.

References

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A, 58*(6), 1065–1086. https://doi.org/10.1080/02724980443000476

Barnes, A. E., Kim, Y.-S., Tighe, E. L., & Vorstius, C. (2017). Readers in adult basic education. *Journal of Learning Disabilities, 50*(2), 180–194. https://doi.org/10.1177/0022219415609187

Binder, K., & Borecki, C. (2008). The use of phonological, orthographic, and contextual information during reading: A comparison of adults who are learning to read and skilled adult readers. *Reading and Writing, 21*(8), 843–858. https://doi.org/10.1007/s11145-007-9099-1

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44*(4), 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 660–675. https://doi.org/10.1037/0096-1523.34.3.660

Cerga-Pashoja, A., Gaete, J., Shishkova, A., & Jordanova, V. (2019). Improving reading in adolescents and adults with high-functioning autism through an assistive technology tool: A cross-over multinational study. *Frontiers in Psychiatry, 10*(546). https://doi.org/10.3389/fpsyt.2019.00546

Cervantes, I.nstituto (2014). Guía del examen C1. https://examenes.cervantes.es/sites/default/files/guia_examen_dele_c1_0.pdf.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 59*(3), 209–217. https://doi.org/10.1037/h0087476

Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L., & Morrison, F. J. (2015). Individual differences in fifth graders' literacy and academic language. *Scientific Studies of Reading, 19*, 114–134. https://doi.org/10.1080/10888438.2014.943905

Dye, L., Hare, D. J., & Hendy, S. (2007). Capacity of people with intellectual disabilities to consent to take part in a research study. *Journal of Applied Research in Intellectual Disabilities, 20*(2), 168–174. https://doi.org/10.1111/j.1468-3148.2006.00310.x

M.inisterio de Educación, C.ultura y Deporte. (2013). *Programa internacional para la evaluación de las competencias de la población adulta. Informe español* http://www.mecd.gob.es/dctm/inee/internacional/piaac/piaac2013vol1.pdf?documentId=0901e72b81741bbc.

A.sociación Española de Normalización. (2018). *Norma Española Experimental UNE 153101 EX. Lectura Fácil: Pautas y recomendaciones para la elaboración de documentos*. AENOR INTERNACIONAL S.A.U.

Everatt, J., & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, and eye movement control. *Language and Speech, 37*(3), 283–297. https://doi.org/10.1177/002383099403700305

Field, A. (2009). *Discovering statistics using SPSS: (and sex, drugs and rock'n'roll)* (3rd ed.). SAGE.

Fracasso, L. E., Bangs, K., & Binder, K. S. (2014). The contributions of phonological and morphological awareness to literacy skills in the Adult Basic Education population. *Journal of Learning Disabilities, 49*(2), 140–151. https://doi.org/10.1177/0022219414538513

Freyhoff, G., Hess, G., Kerr, L., Menzel, E., Tronbacke, B., & Van der Veken, K. (1998). *Make it simple. European guidelines for the production of easy-to-read information for people with learning disability*. ILSMH European Association.

Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. *AERA Open, 6*(1), 1–20. https://doi.org/10.1177/2332858420901696

Hall, R., Greenberg, D., Laures-Gore, J., & Pae, H. K. (2014). The relationship between expressive vocabulary knowledge and reading skills for adult struggling readers. *Journal of Research in Reading, 37*(1), 87–100. https://doi.org/10.1111/j.1467-9817.2012.01537.x

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*(2), 95–123. https://doi.org/10.1006/ceps.1998.0987

Hibbing, A. N., & Rankin-Erickson, J. L. (2003). A picture is worth a thousand words: Using visual images to improve comprehension for middle school struggling readers. *The Reading Teacher, 56*(8), 758–770. https://www.jstor.org/stable/20205292.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. B. T. Deubel (Eds.), *The mind's eye* (pp. 313–334). North-Holland. https://doi.org/10.1016/B978-044451020-4/50018-9.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1430–1440. https://doi.org/10.1037/0278-7393.21.6.1430

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431–439. https://doi.org/10.3758/BF03208203

Jian, Y.-C. (2017). Eye-movement patterns and reader characteristics of students with good and poor performance when reading scientific text with diagrams. *Reading and Writing, 30*(7), 1447–1472. https://doi.org/10.1007/s11145-017-9732-6

Jian, Y.-C. (2018). Reading instructions influence cognitive processes of illustrated text reading not subject perception: An eye-tracking study. *Frontiers in Psychology, 9*, 2263. https://doi.org/10.3389/fpsyg.2018.02263

Jian, Y.-C., & Wu, C. J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text-diagram integration. *Journal of Science Education Technology, 24*(1), 43–55. https://doi.org/10.1007/s10956-014-9519-3

Jones, F. W., Long, K., & Finlay, W. M. L. (2007). Symbols can improve the reading comprehension of adults with learning disabilities. *Journal of Intellectual Disability Research, 51*(7), 545–550. https://doi.org/10.1111/j.1365-2788.2006.00926.x

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Karreman, J., van der Geest, T., & Buursink, E. (2007). Accessible website content guidelines for users with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 20*(6), 510–518. https://doi.org/10.1111/j.1468-3148.2006.00353.x

Kaufman, A. S., Kaufman, N. L., Calonge, I., & y Cordero, A. (2000). In K. BIT (Ed.), *Test breve de inteligencia de Kaufman*. TEA.

Liu, C. J., Kemper, S., & McDowd, J. (2009). The use of illustration to improve older adults' comprehension of health-related information: Is it helpful? *Patient Education and Counseling, 76*(2), 283–288. https://doi.org/10.1016/j.pec.2009.01.013

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

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*(1), 95–109. https://doi.org/10.1016/j.compedu.2012.07.011

Mason, L., Tornatora, M. C., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing, 28*(6), 851–872. https://doi.org/10.1007/s11145-015-9552-5

Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*(1), 64–73. https://doi.org/10.1037/0022-0663.88.1.64

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715–726. https://doi.org/10.1037//0022-0663.82.4.715

McCartney, K., Burchinal, M., & Bub, K. L. (2006). Best practices in quantitative methods for developmentalists. *Monographs of the Society for Research in Child Development, 71*, 1–145. https://doi.org/10.1111/j.1540-5834.2006.07103001.x

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

Morton, J. (1969). Categories of interference: Verbal mediation and conflict in card sorting. *British Journal of Psychology, 60*(3), 329–346. https://doi.org/10.1111/j.2044-8295.1969.tb01204.x

Poncelas, A., & Murphy, G. (2007). Accessible information for people with intellectual disabilities: Do symbols really help? *Journal of Applied Research in Intellectual Disabilities, 20*(5), 466–474. https://doi.org/10.1111/j.1468-3148.2006.00334.x

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191–201. https://doi.org/10.3758/BF03197692

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: eye movements during reading and scanning are different. *Perception & Psychophysics, 58*(5), 734–747. BF03213106. https://doi.org/10.3758/.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1188–1200. https://doi.org/10.1037/0096-1523.22.5.1188

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioural and Brain Sciences, 26* (4), 445-526. https://doi.org/10.1017/s0140525x03000104

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253

Rivero-Contreras, M., Engelhardt, P. E., & Saldaña, D. (2021). An experimental eye-tracking study of text adaptation for readers with dyslexia: Effects of visual support and word frequency. *Annals of Dyslexia*. https://doi.org/10.1007/s11881-021-00217-1

Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 1329–1362). International Reading Association.

Sadoski, M., & Paivio, A. (2012). Dual coding in literacy. In *Imagery and text: A dual coding theory of reading and writing* (pp. 28–48). Taylor & Francis Group.

Sebastián, N. (Ed.). (2000). *LEXESP. Léxico informatizado del español*. Universidad de Barcelona.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1280–1293. https://doi.org/10.1037/a0016896

Taylor, N. A., Greenberg, D., Laures-Gore, J., & Wise, J. C. (2012). Exploring the syntactic skills of struggling adult readers. *Reading and Writing: An Interdisciplinary Journal, 25*(6), 1385–1402. https://doi.org/10.1007/s11145-011-9324-9

Tighe, E. L., & Schatschneider, C. (2016a). A quantile regression approach to understanding the relations among morphological awareness, vocabulary, and reading comprehension in Adult Basic Education students. *Journal of Learning Disabilities, 49*(4), 424–436. https://doi.org/10.1177/0022219414556771

Tighe, E. L., & Schatschneider, C. (2016b). Examining the relationships of component reading skills to reading comprehension in struggling adult readers: A meta-analysis. *Journal of Learning Disabilities, 49*(4), 395–409. https://doi.org/10.1177/0022219414555415

Tronbacke, B. (1997). *Guidelines for easy-to-read materials*. IFLA.

Underwood, G., Hubbard, A., & Wilkinson, H. (1990). Eye fixations predict reading comprehension: The Relationships between reading skill, reading speed, and visual inspection. *Language and Speech, 33*(1), 69–81. https://doi.org/10.1177/002383099003300105

Waddill, P. J., McDaniel, M. A., & Einstein, G. O. (1988). Illustrations as adjuncts to prose: A text-appropriate processing approach. *Journal of Educational Psychology, 80* (4), 457–464. https://doi.org/10.1037/0022-0663.80.4.457

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17*(2), 143–154. https://doi.org/10.1016/S0022-5371(78)90110-X

Zargar, E., Adams, A. M., & Connor, C. M. (2020). The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: An eye-movement study. *Reading and Writing, 33*, 511–545. https://doi.org/10.1007/s11145-019-09966-3