



Universiteit
Leiden
The Netherlands

On the readability of texts presented in sentence-by-sentence segments to beginner readers: evidence from self-paced reading and eye tracking.

Koornneef, A.

Citation

Koornneef, A. (2023). On the readability of texts presented in sentence-by-sentence segments to beginner readers: evidence from self-paced reading and eye tracking. *Reading & Writing*.
doi:10.1007/s11145-023-10423-5

Version: Publisher's Version
License: [Creative Commons CC BY 4.0 license](#)
Downloaded from: <https://hdl.handle.net/1887/3714077>

Note: To cite this publication please use the final published version (if applicable).



On the readability of texts presented in sentence-by-sentence segments to beginner readers: Evidence from self-paced reading and eye tracking

Arnout Koornneef¹ 

Accepted: 1 February 2023
© The Author(s) 2023

Abstract

Many digital reading applications have built-in features to control the presentation flow of texts by segmenting those texts into smaller linguistic units. Whether and how these segmentation techniques affect the readability of texts is largely unknown. With this background, the current study examined a recent proposal that a sentence-by-sentence presentation mode of texts improves reading comprehension of beginning readers because this presentation mode encourages them to engage in more effortful sentence wrap-up processing. In a series of self-paced reading and eye-tracking experiments with primary school pupils as participants (6–9 years old; $n = 134$), reading speed and text comprehension were assessed in a full-page control condition—i.e., texts were presented in their entirety—and in an experimental condition in which texts were presented in sentence-by-sentence segments. The results showed that text comprehension scores were higher for segmented texts than for full-page texts. Furthermore, in the final word-regions of the sentences in the texts, the segmented layout induced longer reading times than the full-page layout did. However, mediation analyses revealed that these inflated reading times had no, or even a disruptive influence on text comprehension. This indicates that the observed comprehension advantage for segmented texts cannot be attributed to more effortful sentence wrap-up. A more general implication of these findings is that the segmentation features of reading applications should be used with caution (e.g., in educational or professional settings) because it is unclear how they affect the perceptual and cognitive mechanisms that underlie reading.

Keywords Readability · Text comprehension · Reading speed · Eye tracking · Segmented layout · Reading applications · Primary school · Reading acquisition

✉ Arnout Koornneef
a.w.koornneef@fsw.leidenuniv.nl

¹ Department of Education and Child Studies, Leiden University, Pieter de La Court Building, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

Introduction

An important goal of research on reading comprehension is to decipher how the readability of texts can support people to increase their focus during reading, thereby assisting them to construct high-quality mental representations of what texts are about. With the rise of computer-based reading applications that can alter the visual appearance of texts on the fly, this endeavor should have become ever more prevalent. However, relatively few studies thoroughly assessed the efficacy and design principles of such applications. In fact, some scholars of reading research tend to disregard reading applications because, in their view, these applications will only have a minor impact on reading processes—or their (unconventional) design features may even disrupt reading (cf. Benedetto et al., 2015; Koornneef & Kraal, 2022; Rayner et al., 2016). In my view the rapid developments in digital text design should receive more attention. They present new (or renewed) windows into applied and fundamental questions on reading processes and may be particularly relevant for the challenges that beginner readers face in becoming proficient readers (cf. Schneps et al., 2019). The simplified layouts and adjusted presentation modes that digital applications offer to users may provide scaffolds for children in primary school to overcome these challenges. The current study contributes to this line of research by exploring potential benefits of a *segmented* layout in which texts are presented in a *sentence-by-sentence* fashion. According to Koornneef et al. (2019), such a simplified, single-sentence reading mode constitutes a very suitable way of presenting texts to beginner readers because it can function as a catalyst to promote more effortful reading. This proposal was addressed in two experiments with young readers in primary school (6–9 years old).

Segmented texts

It is beyond the scope of the current contribution to present an overview of all the segmentation features that are implemented in reading-assistance applications that have been developed over the past few decades. Instead, a brief illustration is provided of how two popular applications, Immersive Reader and Spritz, allow readers to process a text in a segmented fashion. Immersive reader is a feature-rich application that is implemented in major office software packages. In addition to many traditional features to adjust the presentation mode of a text (e.g., letter spacing, line spacing, line length, voice narration), Immersive Reader offers a more experimental feature (line focus) that can be used to channel readers' attention to specific segments of a text by highlighting sets of one, three, or five lines of text. With the application Spritz, a text is displayed in word-by-word segments in the middle of the screen for a fixed duration. As a result, the necessity to plan and execute saccades during reading (i.e., the rapid jumps of the eyes between fixations) is reduced or even eliminated. Hence, in both Immersive Reader and Spritz the flow of presentation can be controlled by segmenting texts into smaller units. According to the developers of such reading applications, this should sustain attention, optimize

word recognition, and streamline oculomotor control processes during reading. As a result, readers can focus their attention on the content of texts, enabling them to construct accurate and elaborate mental representations (see Immersive Reader, 2022; Spritz, 2022).

However, the evidence in support of these speculative hypotheses is scarce. For example, to my knowledge there are no published studies on the efficacy of segmentation techniques such as the line focus feature of Immersive Reader. Furthermore, the segmented RSVP (Rapid Serial Visual Presentation) principles of Spritz have been criticized because the suppression of eye movements may increase cognitive load and visual fatigue (Benedetto et al., 2015; but see Ricciardi & Di Nocera, 2017). In addition, Spritz decreases the accuracy of text comprehension by preventing regressive eye movements that support higher-level linguistic processing (e.g., repairing inconsistencies and misinterpretations) (Schotter et al., 2014; but see Koornneef et al., 2019, for a defense of Spritz). Together this raises questions on whether text segmentation as implemented in popular reading applications is truly a useful technique to assist readers in processing the content of texts.

A methodological study on self-paced reading by Chung-Fat-Yim et al. (2017) bears to this issue. Self-paced reading is a research technique that is frequently used in psycholinguistic studies of language processing. Texts are presented in a chunk-by-chunk (e.g., word-by-word, sentence-by-sentence) manner, with participants asked to press a key or button to progress to the next chunk of text. This procedure allows researchers to study fundamental cognitive processes at the word, syntactic, sentence, and discourse level because they can measure exactly when and for how long certain linguistic content is processed by participants. Chung-Fat-Yim et al. (2017) examined the ecological validity of self-paced reading by comparing sentence-by-sentence reading to a control condition in which a text was presented in a full-page format. They observed that a segmented presentation mode resulted in longer reading times and higher comprehension. According to the authors, an explanation for these findings is that readers in the segmented condition are more inclined to read each sentence carefully, as this is the only opportunity for them to process this information (participants could not re-read prior sentences after a key had been pressed). Hence, they postulate that readers adopt a more effortful processing approach when confronted with segmented texts.

Chung-Fat-Yim et al. (2017) presented this hypothesis as a speculative post-hoc interpretation of their results. A more recent study by Koornneef et al. (2019) with young, beginner readers revealed an identical pattern. In comparison to a control presentation mode with a traditional continuous layout (i.e., texts were presented in their entirety and sentences continued on the same line as far as the width of the text window allowed), word-by-word and sentence-by-sentence presentation modes induced longer reading times and improved comprehension. The authors proposed that in particular sentence-by-sentence reading is useful for beginner readers, based on a similar rationale as put forward by Chung-Fat-Yim et al. (2017): Because the children cannot re-read prior sentences, they are stimulated to process the available visible input more accurately than they would normally do. In fact, Koornneef et al. (2019) speculated that a sentence-by-sentence presentation mode may reflect “an optimal layout for ... second- and third-grade pupils. Because these readers still

heavily rely on sentence-final wrap up to obtain an integrated representation of a text [(Tiffin-Richards & Schroeder, 2018)], stimulating them to do so may increase the quality of their mental representation of a text” (p. 340).

Current study

Text segmentation is potentially a powerful technique to increase the focus of beginner readers, but it is also a controversial issue that requires further testing. The study presented here contributes to this endeavor by examining several aspects of the hypothesis that a segmented, sentence-by-sentence presentation layout induces an effortful reading strategy with improved text comprehension as a result (Chung-Fat-Yim et al., 2017; Koornneef et al., 2019). The participants were young readers in primary school because sentence-by-sentence reading may be particularly useful for beginner readers (Koornneef et al., 2019). There were four aims to the study. The first aim was to replicate prior findings that in comparison to a full-page presentation layout, reading times are longer and comprehension accuracy is improved for a sentence-by-sentence presentation layout. A second aim was to examine whether reading medium (paper vs. screen) moderates the impact of segmented texts on comprehension accuracy because there is some evidence that digital texts induce a more shallow reading strategy than texts presented on paper (e.g., readers are less inclined to construct coherent situation models of digital texts; see Delgado et al., 2018, and Furenes et al., 2021, for meta-analysis studies on ‘the Shallowing Hypothesis’ and ‘screen inferiority’ effects). This raises the possibility that the comprehension advantage for segmented texts is not an inherent property of this layout but emerges in interaction with the factor medium (i.e., text segmentation is only—or primarily—effective for digital texts). The third aim was to evaluate the hypothesis that a sentence-by-sentence presentation layout encourages readers to slow down in the final regions of a sentence, to allocate more cognitive resources to sentence wrap-up processes (Koornneef et al., 2019). Neither Chung-Fat-Yim et al. (2017) nor Koornneef et al. (2019) tested for a causal relationship between decreased reading speed and improved reading comprehension. Therefore, the fourth aim was to show with statistical mediation analyses that reading-time and comprehension effects are not only loosely correlated, but that they are indicative of a mechanism in which segmented texts induce more effortful reading (reflected by longer reading times) and *thereby* improve the mental representation of texts.¹ These four aims were addressed in a self-paced reading experiment (Experiment 1) and an eye-tracking experiment (Experiment 2).

¹ For discussions on the relationship between reading times and reading strategies see, for example, Van Moort et al. (2022) and Yeari et al. (2015).

Experiment 1

Experiment 1 was a replication study of the second experiment in Koornneef et al. (2019). Digital texts were presented in their entirety on a screen in a continuous manner (i.e., sentences continued on the same line as far as text-window width allowed) or they were presented in a segmented manner (sentences were presented one-by-one, forcing readers to process each sentence of a text separately before moving on to the next sentence of the text). This within-participant factor will be referred to as Layout. The reading times for the texts were recorded and each text was followed by a series of questions to assess comprehension accuracy. The design of Koornneef et al. (2019) was extended with the between-participant factor Medium (screen vs. paper): Half of the participants were assigned to an exact replication of Koornneef et al. (2019) and the other half were assigned to a newly developed paper version of that experiment. The following results were predicted: (1) In comparison to full-page texts, segmented texts should induce longer reading times and improved comprehension accuracy; (2) In the case of an interaction effect between Layout and Medium, an increase in comprehension accuracy in the segmented condition should be more prominent in the screen version than in the paper version of the experiment; (3) An increase in comprehension accuracy in the segmented condition should be mediated by reading time.

Materials and methods

Participants

Participants were 80 pupils (39 girls; mean age 8.0 years; range 6.8–9.2) in second ($n=42$) and third grade from five primary schools in The Netherlands. In both experiments reported in the present study, the children had no diagnosed behavioral or attentional problems, and normal or corrected-to-normal vision. The parents or guardians signed a letter of active consent before testing. The children received an eraser after testing.

Texts and comprehension questions

The stimuli consisted of six texts (including two practice texts) that were used in several prior studies (Koornneef & Kraal, 2022; Koornneef et al., 2019; Kraal et al., 2019). The four critical texts consisted of a mix of expository texts (one about the social structure of a community of lions and one about the human skeleton) and narrative texts (one about children who play hide-and-seek at school and one about siblings who encounter a problem with their sister's tablet). The texts consisted of 19 sentences each and the average length was 123 words (range: 117–131 words). An algorithm that calculates text difficulty at the level of conceptual readability in Dutch (P-CLIB version 3.0, Evers, 2008) showed that the texts were age appropriate with scores of CLIB-4 (the equivalent of a text-difficulty level for second grade)

for narrative texts and slightly higher scores of CLIB-5 (the equivalent of a text-difficulty level for third grade) for expository texts. To assess text comprehension, six open-ended questions of different types were posed after each text (i.e., questions tapping literal information, text-based questions requiring a text-connecting inference, and knowledge-based questions requiring a more elaborate inference). The answers of the children were scored as correct or incorrect—based on a strict coding protocol containing exhaustive lists of examples of correct and incorrect answers. For more detailed information about the materials, see Kraal et al. (2019).

Design and procedure

All procedures were approved by the Leiden University Institute of Education and Child Studies ethics committee (project number ECPW-2015/107) and conducted in accordance with the Declaration of Helsinki. Participants were tested individually in a quiet room. The duration of a test session was 20–30 min. There were two versions of the experiment: A *screen* version ($n=40$) and a *paper* version ($n=40$).

The open-source server Ibex Farm (Drummond, 2013) and its supplementary software were used to run the *screen* version of the experiment on a laptop or desktop computer at the schools of the participants. The experiment consisted of two main blocks. Both blocks started with instructions and a practice text to familiarize the participants with the procedures. The participants were not allowed to use finger tracking techniques (i.e., the movement of a child's index finger that points to printed text while reading) to support the reading process—neither in the screen version nor in the paper version of the study. The practice phase of a block was followed by a testing phase in which the children read two texts for comprehension (one narrative text, one expository text). In one block, texts were presented in their entirety (*full-page* layout). The children pressed the space bar to make a text appear on the computer screen. After completion of the text, they pressed the space bar another time to progress to the comprehension questions. The elapsed time between space-bar presses was recorded to obtain total reading times for the texts. In the other block, each sentence of a text was presented separately (*segmented* layout). The children pressed the space bar to make the first sentence of a text appear. After pressing the space bar again, the first sentence of the text was replaced by its second sentence. By repeatedly pressing the space bar the child read all the sentences of a text. It was not possible to go back to sentences presented earlier in the trial. The elapsed time between space-bar presses was recorded to obtain the reading times of a text. Six comprehension questions appeared on screen one by one after each text. The test leader read out aloud the question and recorded the responses. The ordering of the two experimental blocks and the four critical texts was rotated across four counterbalanced lists. Participants were assigned to one of those lists.

The *paper* version of Experiment 1 was created by taking high-resolution screenshots of each digital page of the screen version of the experiment. These pages were printed on paper and bundled in four booklets (one for each counterbalanced list). Participants were assigned to one of those booklets. The instructions and procedures were identical to the screen version, yet, instead of pressing the space bar to progress

Table 1 Mean (*M*) accuracy scores (probability correct), mean reading times (in milliseconds per word), and their standard deviations (*SD*) as a function of Medium and Layout (Experiment 1)

Layout	Paper		Screen			
	Accuracy		Accuracy		Reading times	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Full-page	0.57	0.25	0.60	0.25	631	265
Segmented	0.65	0.23	0.63	0.20	731	269

through the experiment, the children flipped the pages in the booklet. Reading times were not recorded in the paper version of the experiment.

Results

Descriptive results are reported in Table 1. The reading times reflect the average reading time (in milliseconds) of the words in a text (i.e., the reading time for a text was divided by the number of words in that text). For the statistical analyses, the reading times per word were log transformed to correct for right skew. Mixed-effects logistic regression models were fitted for the comprehension questions and mixed-effects linear regression models were fitted for the transformed reading times. The models were fitted with R (R Core Team, 2021) using the package *lme4* (Bates et al., 2015). Wald chi-square testing (Type II) as implemented in the package *car* (Fox & Weisberg, 2019) was applied to test for main and interaction effects. Follow-up analyses were conducted with the package *emmeans* (Lenth, 2021). The *p*-values in the follow-up analyses were based on asymptotic degrees of freedom (i.e., *z*-statistics). Tables for the models were created with *sjPlot* (Lüdtke, 2021) and figures were plotted with *ggplot2* (Wickham, 2016).

Accuracy

The statistical model included the fixed effects Layout and Medium, as well as their interaction. Participants ($n=80$) and items ($n=24$) were included as crossed random effects (see Table 2 and Fig. 1). Wald chi-square tests revealed a main effect of Layout ($\chi^2(1)=5.84$, $p=.016$) indicating that the children performed better on the comprehension questions in the segmented condition than in the full-page condition ($\hat{\beta}=0.27$, $SE=0.11$). Neither a main effect of Medium ($\chi^2(1)=0.05$, $p=.831$) nor an interaction effect of Layout and Medium ($\chi^2(1)=0.66$, $p=.418$) was observed.

Reading times

The model included the fixed effect Layout (Medium was not included as reading times were not recorded in the paper condition). Participants ($n=40$) and items ($n=4$) were included as crossed random effects (see Table 3 and Fig. 1). Wald chi-square tests revealed an effect of Layout ($\chi^2(1)=28.31$, $p<.001$) indicating

Table 2 Summary of the Medium * Layout mixed-effects model for the accuracy scores (Experiment 1)

Predictors	Accuracy		
	Log-Odds	SE	CI
(Intercept)	0.45	0.29	-0.12–1.01
Medium [Screen]	0.14	0.27	-0.40–0.68
Layout [Segmented]	0.36	0.16	0.05–0.67
Medium [Screen] * Layout [Segmented]	-0.18	0.22	-0.61–0.25
<i>Random effects</i>			
τ_{00} participants	1.02		
τ_{00} items	1.10		
<i>N</i> participants	80		
<i>N</i> items	24		
Observations	1920		

The reference level in the model is the paper/full-page condition. *SE*=standard error; *CI*=95% confidence interval; τ_{00} =variance of random effects

Table 3 Summary of the mixed-effects model for the log-transformed reading times in the screen condition (Experiment 1)

Predictors	Log reading times		
	Estimates	SE	CI
(Intercept)	6.37	0.06	6.26–6.48
Layout [Segmented]	0.17	0.03	0.11–0.23
<i>Random effects</i>			
τ_{00} participants	0.10		
τ_{00} items	0.00		
<i>N</i> participants	40		
<i>N</i> items	4		
Observations	160		

Full-page condition is the reference level

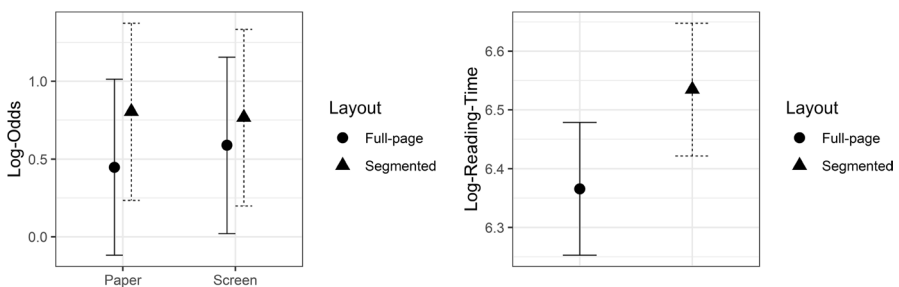


Fig. 1 Fixed effects estimates and their 95% confidence intervals of the accuracy scores (probability correct; left figure) and log-transformed reading times (in milliseconds per word; right figure) as a function of Medium and Layout (Experiment 1)

that children read more slowly in the segmented condition than in the full-page condition.

Mediation

Mediation analyses were performed with the R-package *bmlm* (Bayesian Multi-Level Mediation; Vuorre, 2017), allowing 1–1–1 (lower-level) mediation for data with a multi-level structure. The *isolate*-function of *bmlm* was used to compute within-subject text-by-text deviations from the subject means representing a within-person version of the log reading times (Vuorre & Bolger, 2017). The model was fitted with the dummy-coded variable Layout (0 = full-page; 1 = segmented) as independent variable (IV), the binary variable accuracy (0 = incorrect; 1 = correct) as dependent variable (DV), and within-person log reading times as mediator (M). The default settings of *bmlm* were applied—which are carefully selected to have minimal impact on the resulting posterior distributions, given common ranges of data values (Vuorre & Bolger, 2017). The results revealed no credible mediation effect ($\hat{\beta} \approx 0.00$, $SE = 0.16$; 95% credible interval $[-0.33-0.29]$).

Summary of results Experiment 1

Consistent with the first prediction, the analyses showed that segmented texts are read more slowly than full-page texts and that the comprehension scores for segmented texts are higher than for full-page texts. Concerning the second prediction, the positive effect of text segmentation on comprehension accuracy was of a similar magnitude in digital and paper texts. Contrary to the third prediction, the results for reading times and reading comprehension could not be explained by a mechanism in which increased processing times for segmented texts are the underlying cause of improved comprehension. The findings of Experiment 1 are discussed further in the "[Discussion](#)" section.

Experiment 2

Koornneef et al. (2019) postulated that the main advantage of a sentence-by-sentence presentation mode is that readers slow down near the end of each sentence, right before they press a button or key to proceed to the next sentence. This slow-down in reading allows readers to engage in more elaborate sentence wrap-up, resulting in better mental representations of texts. The findings of Experiment 1 are inconsistent with this hypothesis as no mediation effect of reading time was observed. However, the self-paced reading method that was used does not provide a proper test as it only measures reading times at the sentence and text level. To address this limitation, eye tracking was used in Experiment 2. This methodology provides a wealth of data about reading behavior by monitoring readers' eye gaze from millisecond to millisecond. As such it can provide detailed information about *where* in a sentence or text readers slow down (or speed up) and thus offers a more

careful examination of the hypothesis that sentence-by-sentence reading stimulates more elaborate sentence wrap-up.

The same design was used as in the screen version of Experiment 1: The participants read both full-page and segmented texts on a computer screen while their eye movements were recorded. The following results were predicted: (1) In comparison to full-page texts, segmented texts should induce longer reading times and improved comprehension accuracy; (2) Slower reading of segmented texts should emerge in the final word regions of the sentences of the texts; (3) An increase in comprehension accuracy in the segmented condition should be mediated by the reading times of these final regions of sentences.

Materials and methods

Participants

Participants were 54 pupils (30 girls; mean age 8.2 years; range 7.1–9.7) in second ($n=35$) and third grade from two primary schools in The Netherlands. None of them participated in Experiment 1.

Texts and comprehension questions

The same texts and comprehension questions were presented as in Experiment 1.

Design and procedure

The design and procedures were kept identical to the screen version of Experiment 1 as much as possible. An Eyelink 1000 setup (SR Research) and its supplementary software were used to record (Experiment Builder) and pre-process (Data Viewer) the children's eye movements. The stimuli were displayed about 60 cm from the participants' eyes on a computer screen. The children rested their head on a chin-rest to prevent them from moving their head. In the full-page condition, texts were presented with increased spacing between lines to improve data pre-processing decisions—e.g., fixations could be assigned to the correct word of a line, even if mild tracker loss or drifting occurred. Before the children read each text, their right eye was calibrated using nine fixation points. Each trial began with a single fixation point presented just to the left of the first character of the upcoming text or sentence—in the segmented condition this fixation point was presented before each sentence. After completion of each text, the children moved their head from the chin-rest to answer the comprehension questions. The duration of a test session was 30–50 min.

Results

For the analyses the same approach and R-packages were used as for Experiment 1.

Table 4 Summary of the mixed-effects model for the accuracy scores (Experiment 2)

<i>Predictors</i>	<i>Accuracy</i>		
	<i>Log-Odds</i>	<i>SE</i>	<i>CI</i>
(Intercept)	0.89	0.28	0.34–1.44
Layout [Segmented]	0.24	0.14	−0.03–0.51
<i>Random effects</i>			
τ_{00} participants	0.58		
τ_{00} items	1.37		
<i>N</i> participants	54		
<i>N</i> items	24		
Observations	1289		

Full-page condition is the reference level

Accuracy

Seven answers to the comprehension questions (0.5% of the data) were not recorded properly and removed from the analyses. The model for the remaining data on accuracy included the fixed effect Layout, the random effect participants ($n=54$), and the random effect items ($n=24$) (see Table 4). Wald chi-square tests revealed no reliable effect of Layout ($\chi^2(1)=3.13$, $p=.077$) but the direction of the effect was as predicted, with better performance on the comprehension questions in the segmented condition ($M=0.70$, $SD=0.23$) than in the full-page condition ($M=0.65$, $SD=0.26$).

Reading times

Due to tracker loss, the recordings of four trials (1.9% of the data) were removed from the analyses. The words of a text were treated as separate areas of interest (AOIs), indicated by rectangular shapes that were drawn around each individual word by an automatic procedure of Eyelink's software package Data Viewer. Subsequently, the total reading time for each word in each text was computed for every participant. This measure was computed by taking the sum of all fixation durations of that word, including re-fixations and fixations that were made on the word by regressive eye movements. The recordings for words that were not fixated ('skipped') during reading were treated as missing data. As a final preparational step, the words within each sentence were coded as *first* (first word of a sentence), *second* (second word of a sentence), *prefinal* (prefinal word of a sentence), *final* (final word of a sentence), and *mid* (for all remaining words in a sentence) (see Table 5 for descriptive values).

The mixed-effects model for the log-transformed reading times included fixed effects of Layout, Region, and their interaction. Participants ($n=54$) and items (i.e., the unique sentences in the experiment, $n=76$) were included as crossed random effects (see Table 6). Wald chi-square tests revealed effects of Layout ($\chi^2(1)=44.25$,

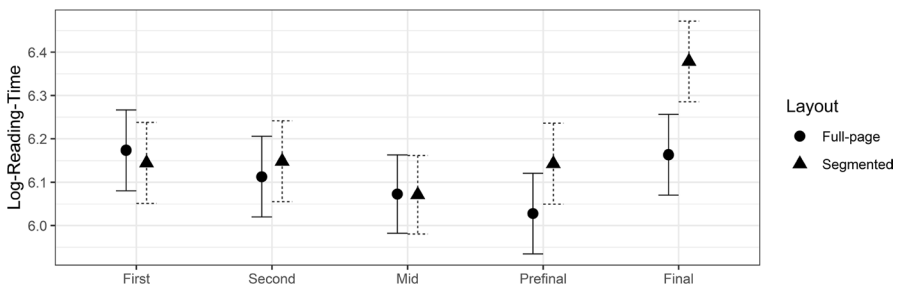
Table 5 Mean reading times (in milliseconds) and their *SDs* as a function of Layout and Region (Experiment 2)

Layout	Region									
	First		Second		Mid		Prefinal		Final	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Full-page	607	324	598	342	542	255	534	238	630	304
Segmented	604	365	600	311	549	294	579	203	772	374

Table 6 Summary of the Layout * Sentence Region mixed-effects model for the log-transformed reading times (Experiment 2)

Predictors	Log reading times		
	Estimates	SE	CI
(Intercept)	6.17	0.05	6.08–6.27
Layout [Segmented]	–0.03	0.02	–0.07–0.01
Region [Second]	–0.06	0.02	–0.10–0.02
Region [Mid]	–0.10	0.02	–0.13–0.07
Region [Prefinal]	–0.15	0.02	–0.18–0.11
Region [Final]	–0.01	0.02	–0.05–0.03
Layout [Segmented] * Region [Second]	0.06	0.03	0.01–0.12
Layout [Segmented] * Region [Mid]	0.03	0.02	–0.02–0.07
Layout [Segmented] * Region [Prefinal]	0.14	0.03	0.09–0.20
Layout [Segmented] * Region [Final]	0.24	0.03	0.19–0.30
<i>Random effects</i>			
τ_{00} participants	0.10		
τ_{00} items	0.01		
<i>N</i> participants	54		
<i>N</i> items	76		
Observations	23,760		

Full-page/First is the reference level

**Fig. 2** Fixed effects estimates and their 95% confidence intervals of the log-transformed reading times as a function of Layout and Region (Experiment 2)

$p < .001$), Region ($\chi^2(4) = 300.05$, $p < .001$), and a Layout X Region interaction ($\chi^2(4) = 114.76$, $p < .001$) (see Fig. 2). Follow-up analyses showed that the children read more slowly in the segmented condition than in the full-page condition at the prefinal ($\hat{\beta} = 0.12$, $SE = 0.02$, $z = 5.79$, $p < .001$) and final ($\hat{\beta} = 0.22$, $SE = 0.02$, $z = 10.96$, $p < .001$) words of a sentence. Other sentence regions showed no effect of Layout (first: $\hat{\beta} = 0.03$, $SE = 0.02$, $z = 1.47$, $p = .141$; second: $\hat{\beta} = 0.04$, $SE = 0.02$, $z = 1.83$, $p = .067$; mid: $\hat{\beta} = 0.00$, $SE = 0.01$, $z = 0.11$, $p = .909$). Furthermore, relative to the words in the middle of a sentence (mid-region), the first, second, and final words of a sentence were read more slowly in the full-page condition (first: $\hat{\beta} = 0.10$, $SE = 0.02$, $z = 6.07$, $p < .001$; second: $\hat{\beta} = 0.04$, $SE = 0.02$, $z = 2.41$, $p = .016$; final: $\hat{\beta} = 0.09$, $SE = 0.02$, $z = 5.52$, $p < .001$) and in the segmented condition (first: $\hat{\beta} = 0.07$, $SE = 0.02$, $z = 4.35$, $p < .001$; second: $\hat{\beta} = 0.08$, $SE = 0.02$, $z = 4.62$, $p < .001$; final $\hat{\beta} = 0.31$, $SE = 0.02$, $z = 18.16$, $p < .001$). A different pattern was observed for the prefinal word. In the full-page condition the prefinal word was read more quickly than words in the mid-region ($\hat{\beta} = 0.05$, $SE = 0.02$, $z = 2.69$, $p = .007$), yet in the segmented condition the prefinal word was read more slowly than words in the mid-region ($\hat{\beta} = 0.07$, $SE = 0.02$, $z = 4.24$, $p < .001$).²

Mediation

Mediation analyses were carried out for the two sentence regions that revealed a significant effect of Layout on the reading times, i.e., the prefinal and final word regions. The results revealed no credible mediation effect at the prefinal region ($\hat{\beta} = 0.03$, $SE = 0.09$; 95% credible interval $[-0.15-0.23]$) but (weak) evidence of a mediation effect was observed for the final word region ($\hat{\beta} = -0.22$, $SE = 0.15$; 95% credible interval $[-0.54-0.04]$): Segmented texts induced longer reading times for the final word of a sentence than full-page texts did (path a; see Fig. 3), but these longer processing times had a disruptive influence on the accuracy scores of readers (path b).

Summary of results Experiment 2

Consistent with the first prediction, segmented texts induced longer reading times than full-page texts did. For the outcome measure comprehension accuracy the direction of the effect was also as predicted—i.e., better performance on the comprehension questions in the segmented condition—but this effect fell short of significance. Consistent with the second prediction, relative to sentences in full-page texts readers slowed down in the final word regions of sentences in segmented texts. This finding appears to be compatible with the idea that sentence-by-sentence reading induces more elaborate wrap-up processes. However, the mediation analyses—which provide a more direct test of this third prediction—revealed an

² Control analyses were conducted in which skipped words were not treated as missing data but received a valid total reading time of zero. These analyses confirmed the results as reported here.

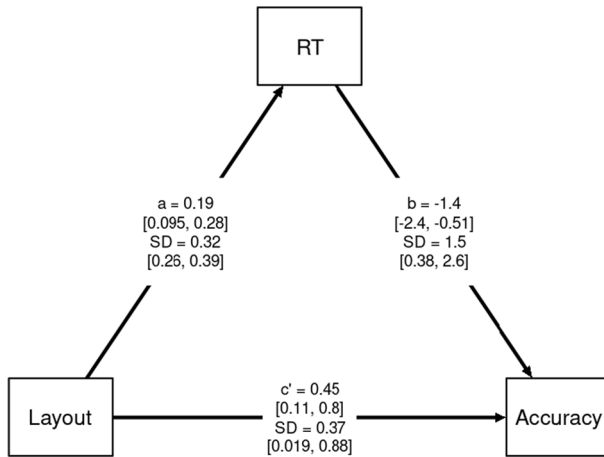


Fig. 3 Path diagram of the mediation analyses ($a * b$ =mediation effect; c' =direct effect) of the final word region in Experiment 2 with point estimates (posterior means) of the parameters and associated 95 percent credible intervals in square brackets below the point estimates (SD =the associated effect's standard deviation indicating the degree to which that effect varies between people; IV =Layout; DV =Accuracy; M =RT)

unanticipated relationship between reading speed and reading comprehension: Longer reading times in the final word region of segmented texts had a disruptive influence on comprehension accuracy. This indicates that a comprehension advantage for segmented texts does not emerge because readers slow down near the end of each sentence, but rather, that this advantage emerges in spite of this slowdown in reading speed. Next to these main findings, it was observed that the processing signatures within a sentence were quite similar across segmented and full-page conditions. That is, in both text formats the reading times for the first, second, and final words of a sentence were longer than for the words in the middle of a sentence. The findings of Experiment 2 are discussed further in the Discussion section below.

Discussion

There were four aims to the current study: (1) To replicate the findings of prior studies that in comparison to a full-page presentation mode, reading times are longer and comprehension accuracy is improved for a sentence-by-sentence presentation mode; (2) To control for the possibility that reading medium (paper vs. screen) moderates the impact of sentence-by-sentence texts on comprehension accuracy; (3) To test the hypothesis that a sentence-by-sentence presentation mode encourages readers to slow down in the final regions of a sentence; (4) To test the hypothesis that these inflated processing times are the underlying cause of improved comprehension.

Regarding Aims 1 and 2, the data of both experiments showed the same pattern as prior studies did, displaying significantly longer reading times and improved comprehension accuracy for a sentence-by-sentence presentation mode, relative to its full-page control presentation mode (Chung-Fat-Yim et al., 2017; Koornneef et al., 2019) (Aim 1). Furthermore, the positive influence of a sentence-by-sentence presentation mode on comprehension accuracy did not interact with type of medium (Aim 2). It should be noted that the current results for comprehension accuracy did not fully replicate prior results as no statistically significant differences between segmented and full-page texts emerged in Experiment 2—but the direction of the effect was as predicted with numerically higher comprehension scores for sentence-by-sentence reading. Although it is not clear why Experiment 2 showed no robust comprehension advantage for segmented texts, two methodological factors should be considered. First, the absence of an effect could be due to a smaller sample size in Experiment 2 ($n=54$) in comparison to Experiment 1 ($n=80$) and the Koornneef et al. (2019) study ($n=88$). Second, in contrast to prior studies, the full-page texts in Experiment 2 were presented with increased spacing between lines. Although this design decision enables easier and more reliable pre-processing steps of the eye-tracking data, it may also have a side effect on the comprehension scores of readers, because line spacing is known to affect the readability of a text. With increased line spacing it will be relatively easy for beginner readers to keep their focus on the line that they are reading and, in addition, return-sweep targets (i.e., the first word on a line of text) are detected with less effort (Madhavan et al., 2016; Vanderschantz, 2008). As a result, comprehension for full-page texts may improve, thereby diminishing the observed comprehension advantage effect for sentence-by-sentence texts in Experiment 2.

Concerning Aims 3 and 4, in comparison to full-page control texts, beginner readers slowed down solely in the final regions of sentences (i.e., at the prefinal and final words) of texts that were presented in segments (Aim 3). The most important hypothesis investigated in the current study was that these increased reading times at the end of a sentence are indicative of a process in which a sentence-by-sentence presentation mode induces more elaborate sentence wrap-up. This could assist readers to create a more interconnected, coherent situation model of a text, allowing them to formulate more accurate answers to the comprehension questions posed after each text. The mediation analyses did not provide support for this hypothesis (Aim 4). In Experiment 1, no mediation effect of reading time on reading comprehension emerged at all, and the analyses of Experiment 2 revealed a (weak) mediation effect in an unanticipated direction. More specifically, longer reading times in the final word region of segmented texts seem to have a disruptive influence on comprehension accuracy, suggesting that a comprehension advantage for segmented texts does not emerge because readers slow down near the end of each sentence, but rather, that this advantage emerges in spite of this slowdown in reading speed. This could imply that increased reading times in segmented texts do not reflect more elaborate wrap-up but index dual-task processing costs instead. After all, the children are required to press the space bar after reading each sentence of the segmented texts, and preparing and executing this additional manual task may interfere with the main task of reading.

If this latter interpretation of the mediation analyses is correct, the observation that there is still an overall (or direct) positive influence of segmented texts on comprehension accuracy becomes somewhat puzzling. Although there is no unequivocal way to solve this puzzle, several scenarios should be considered. One possible scenario is that the inflated reading times at the end of each sentence in segmented texts index two (independent) processes: (1) A process of more elaborate sentence wrap-up; (2) A process of dual-task interference. In this scenario, the interfering process would be more pronounced in the current study, fully masking any mediation effects of readers' sentence wrap-up processing efforts. Another possible scenario is that even though sentence wrap-up processes elicit detectable costs, the amount of time spent on the final words of a sentence does not reflect the quality or effectiveness of those processes. In that case, increased reading time durations will not show a positive correlation with improved performance on the comprehension questions. In these two scenarios, the general hypothesis that sentence-by-sentence reading induces more sophisticated sentence wrap-up processes can be maintained. Yet, it is equally plausible that this general framework of improved sentence wrap-up simply does not hold, and should be discarded. A finding that seems to corroborate this third scenario is that even in full-page texts, children slowed down considerably at the final word of a sentence in comparison to mid-sentence word regions, suggesting that standard interpunctuation in full-page texts may be sufficient to trigger adequate wrap-up processes.

The discussion above reveals that any comprehension advantages for sentence-by-sentence texts could—and perhaps should—be attributed to other characteristics of this presentation mode. For example, return-sweep saccades may be less demanding in a sentence-by-sentence layout than in a traditional full-page layout because readers only shift their eye gaze in a horizontal plane, not in a vertical plane (Koornneef et al., 2019). In addition, the sentence-by-sentence layout in the current study avoids that clausal units are interrupted by a line break, thereby limiting parsing problems for beginner readers during a return sweep (Levasseur et al., 2006; Raban, 1982). Moreover, limiting the possibilities to look back to earlier sections of a text may be beneficial for beginner readers because it prohibits carrying out redundant and excessively long regressive eye movements during backtracking (cf. Schneps et al., 2010)—but note that this is a very controversial issue (see Schotter et al., 2014). Furthermore, although the sentence-by-sentence condition in the current study required the participants to carry out a dual task by pressing a key after reading each sentence of a text, these key presses may also function as 'tactile-kinesthetic' reinforcement. This type of reinforcement may increase comprehension as it provides additional sensory input and punctuates each phrase during reading (Schneps et al., 2010).³ As a final example, comprehension advantages for segmented texts may arise due to a novelty effect. Because most children will be unfamiliar with the procedures of sentence-by-sentence reading, their level of

³ In fact, as suggested by an anonymous reviewer, juggling multiple tasks may even boost readers' comprehension and self-monitoring by activating frontal areas of the brain which are part of an executive control network (e.g., Miyake et al., 2000).

engagement and motivation for the task at hand may be enhanced, resulting in higher comprehension scores.

In all, the current study has several implications for prior research on segmented presentation modes, and in a more general sense, for reading applications that include text segmentation as a feature to improve the readability of texts. First of all, even though the results do not allow a detailed, unambiguous depiction of the mechanisms that underlie the comprehension advantage effect as observed in sentence-by-sentence presentation modes, they clearly show that the proposal by Koornneef et al. (2019)—who attribute the comprehension advantage to more effortful reading—is overly simplistic. Furthermore, the current results emphasize that future studies on the readability of texts should not only include a variety of (dependent) variables to examine different aspects of the readability of texts, but that they should also carry out (mediation) analyses to examine the relationships between these variables. These issues also bear to the presumed advantages of digital reading applications that allow users to control the flow of presentation by segmenting texts into smaller units (e.g., Spritz, Immersive Reader, Spreeder). That is, even though text segmentation may improve comprehension for some readers in some situations, the underlying cognitive mechanisms are poorly understood at best (cf. Koornneef & Kraal, 2022, for a similar conclusion on the use of the application BeeLine Reader). Hence, one should be cautious with using or introducing these types of applications as interventions or scaffolding techniques for beginner readers to help them to overcome the challenges they encounter during reading acquisition in primary school.

Acknowledgements I am grateful to all bachelor and master students who participated as experiment leaders in the test sessions at the schools. Furthermore, I thank the pupils, schools, and parents/guardians for their involvement in this research.

Author contribution Arnout Koornneef: conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft & review & editing, visualization, supervision, project administration.

Declarations

Conflict of interest This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author is not affiliated in any way to the commercial reading applications mentioned in the article. All procedures were approved by the Leiden University Institute of Education and Child Studies ethics committee (project number ECPW-2015/107) and conducted in accordance with the Declaration of Helsinki. The parents or guardians of the participants signed a letter of active consent before testing.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benedetto, S., Carbone, A., Pedrotti, M., Le Fevre, K., Bey, L. A. Y., & Baccino, T. (2015). Rapid serial visual presentation in reading: The case of spritz. *Computers in Human Behavior*, 45, 352–358. <https://doi.org/10.1016/j.chb.2014.12.043>
- Chung-Fat-Yim, A., Peterson, J. B., & Mar, R. A. (2017). Validating self-paced sentence-by-sentence reading: Story comprehension, recall, and narrative transportation. *Reading and Writing*, 30(4), 857–869. <https://doi.org/10.1007/s11145-016-9704-2>
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25, 23–38. <https://doi.org/10.1016/j.edurev.2018.09.003>
- Drummond, A. (2013). *Ibex farm*. Online Server (Software). Retrieved from <https://adrummond.net/ibexfarm>
- Evers, G. (2008). Programma voor berekening Cito LeesIndex voor het basisonderwijs. *P-CLIB versie 3.0* (pp. 1994–2008). Cito: Arnhem.
- Fox, J., & Weisberg, S. (2019). *An R companion to applied Regression (Third)*. Thousand Oaks CA: Sage.
- Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: A meta-analysis. *Review of Educational Research*, 91(4), 483–517. <https://doi.org/10.3102/0034654321998>
- Koornneef, A., & Kraal, A. (2022). Does BeeLine reader's gradient-coloured font improve the readability of digital texts for beginning readers? *Computers in Human Behavior Reports*, 6, 100197. <https://doi.org/10.1016/J.CHBR.2022.100197>
- Koornneef, A., Kraal, A., & Danel, M. (2019). Beginning readers might benefit from digital texts presented in a sentence-by-sentence fashion. But why? *Computers in Human Behavior*, 92, 328–343. <https://doi.org/10.1016/J.CHB.2018.10.024>
- Kraal, A., van den Broek, P. W., Koornneef, A. W., Ganushchak, L. Y., & Saab, N. (2019). Differences in text processing by low- and high-comprehending beginning readers of expository and narrative texts: Evidence from eye movements. *Learning and Individual Differences*, 74, 101752. <https://doi.org/10.1016/j.lindif.2019.101752>
- Lenth, R. V. (2021). *Emmeans: Estimated marginal means, aka least-squares means*. Retrieved from <https://cran.r-project.org/package=emmeans>.
- Levasseur, V. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, 27(3), 423–445. <https://doi.org/10.1017/S0142716406060346>
- Lüdtke, D. (2021). *sjPlot: Data visualization for statistics in social science*. Retrieved from <https://cran.r-project.org/package=sjPlot>.
- Madhavan, I., Izzuddin Hairol, M., & Mohammed, Z. (2016). Spacing improves reading in dyslexic children. *Asia Pacific Journal of Developmental Differences*, 3(1), 3–20. <https://doi.org/10.3850/S2345734114000183>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Raban, B. (1982). Text display effects on the fluency of young readers. *Journal of Research in Reading*, 5(1), 7–28. <https://doi.org/10.1111/j.1467-9817.1982.tb00126.x>
- Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So much to read, so little time. *Psychological Science in the Public Interest*, 17(1), 4–34. <https://doi.org/10.1177/1529100615623267>
- Immersive Reader (2022). *Immersive Reader: Microsoft learning tools*. <https://www.onenote.com/learningtools>
- Ricciardi, O., & Di-Nocera, F. (2017). Not so fast: A reply to Benedetto et al. (2015). *Computers in Human Behavior*, 69, 381–385. <https://doi.org/10.1016/j.chb.2016.12.047>

- Schneps, M. H., Chen, C., Pomplun, M., Wang, J., Crosby, A. D., & Kent, K. (2019). Pushing the speed of assistive technologies for reading. *Mind, Brain, and Education*, 13(1), 14–29. <https://doi.org/10.1111/MBE.12180>
- Schneps, M. H., O’Keeffe, J. K., Heffner-Wong, A., & Sonnert, G. (2010). Using technology to support STEM reading. *Journal of Special Education Technology*, 25(3), 21–33. <https://doi.org/10.1177/016264341002500304>
- Spritz (202214). *Reading reimagined: The worlds best speed reading app*. <https://spritz.com/>
- Tiffin-Richards, S. P., & Schroeder, S. (2018). The development of wrap-up processes in text reading: A study of children’s eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(7), 1051–1063. <https://doi.org/10.1037/xlm0000506>
- Van Moort, M. L., Koornneef, A., Wilderjans, T. F., & Van den Broek, P. (2022). Validation processes and reading purpose: Is validation against knowledge and prior text influenced by reading goal? *Journal of Educational Psychology*, 114(7), 1533–1552. <https://doi.org/10.1037/edu0000769>
- Vanderschantz, N. (2008). Reviewing the understanding of the effects of spacing on children’s eye movements for on-screen reading. In *Computer science working papers*. Department of computer science, University of Waikato. Retrieved from <https://hdl.handle.net/10289/802>.
- Voorre, M. (2017). *Bmlm: Bayesian multilevel mediation*. Retrieved from <https://cran.r-project.org/package=bmlm>
- Voorre, M., & Bolger, N. (2017). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods*, 50(5), 2125–2143. <https://doi.org/10.3758/S13428-017-0980-9>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. New York: Springer.
- Yeari, M., Van den Broek, P., & Oudega, M. (2015). Processing and memory of central versus peripheral information as a function of reading goals: Evidence from eye-movements. *Reading and Writing*, 28(8), 1071–1097. <https://doi.org/10.1007/s11145-015-9561-4>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.