# Does BeeLine Reader's gradient-coloured font improve the readability of digital texts for beginning readers?

Koornneef, A.W.; Kraal, A.

Computers in Human Behavior Reports 6 (2022) 100197

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

# Does BeeLine Reader's gradient-coloured font improve the readability of digital texts for beginning readers?

Arnout Koornneef[a,1,*], Astrid Kraal[b,1]

[a] Department of Education and Child Studies, Leiden University, Pieter de la Court Building, Wassenaarseweg 52, 2333AK, Leiden, the Netherlands
[b] Stichting CED-Groep, Hoofdweg 70, 3067GH, Rotterdam, the Netherlands

A R T I C L E I N F O

*Keywords:*
Reading acquisition
Primary education
Readability
Font colour
Text layout
BeeLine Reader
Text comprehension
Reading speed

A B S T R A C T

In two reading experiments, we examined the efficacy of the commercial reading assistance application BeeLine Reader which colours the letters of digital texts in gradients. According to its developers, BeeLine Reader increases reading speed, improves comprehension, and makes reading more enjoyable. We tested these hypotheses for second- and third-grade pupils (6–9 years old), assessing the influence of BeeLine Reader in several layouts in which we varied other features that are known to impact the reading processes of beginning readers (line spacing, line length, text segmentation). In comparison to control texts with a standard black font, reading time advantages for BeeLine texts emerged for pupils in second grade (not in third grade) when they read texts with long lines and little inter-line spacing. However, when second-grade readers processed texts that were optimized for their reading level (texts with short lines and sufficient inter-line spacing) they displayed a slower reading pace in texts with a BeeLine font than in texts with a black font. Furthermore, BeeLine texts may hamper comprehension for third-grade readers and were rated as more difficult and less convenient to process than texts with a black font. In conclusion, the visual anchors offered by BeeLine Reader may be useful for some beginning readers in some situations but the application can also impede the readability of texts. These findings emphasize that claims made for digital reading applications should be formally tested if they are going to be introduced into educational settings.

## 1. Introduction

There are numerous digital tools to improve the readability of a text by modifying its visual appearance or presentation mode. Although these reading applications differ in their approach to optimize the reading routines of people, they do share a common ground: they are designed to enhance visual word recognition and oculomotor control by reducing visual distraction and increasing the reader's focus on what is being read. Perhaps the best-known example in the academic literature is Spritz (e.g., Benedetto et al., 2015; Ricciardi & Di Nocera, 2017) which offers the possibility to present a digital text in a word-by-word fashion, reducing (or even eliminating) the necessity to plan and execute saccades during reading (the rapid jumps of the eyes between fixations). Another example – and the main topic of the current study – is BeeLine Reader. BeeLine Reader is a charged application that adds colour gradients to digital texts (e.g., websites, pdf files), yet maintains the overall layout (e.g., line spacing, line length, font type) of the original texts (see Fig. 1). According to its developers, the colour modifications generated by BeeLine Reader should allow for smooth transitions across lines of text (return sweeps).

Scholars of reading research tend to criticize – or in some cases even ignore – the design features of digital reading applications such as BeeLine Reader. In their view, these applications violate well-known principles of reading and developers often fail to put their applications to the empirical test (see e.g., Benedetto et al., 2015; Rayner et al., 2016). This line of criticism is accurate, yet, in our opinion the rapid developments in digital text design should receive more attention. First, some reading assistance applications are already implemented in major software packages and should be studied more systematically. Second, the design features of many digital reading applications raise crucial questions on fundamental

* Corresponding author.
*E-mail addresses:* a.w.koornneef@fsw.leidenuniv.nl (A. Koornneef), a.kraal@cedgroep.nl (A. Kraal).

[1] We are grateful to all bachelor and master students who participated as experiment leaders in the test sessions at the schools. Furthermore, we thank the pupils, schools, and parents/guardians for their involvement in this research. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors are not affiliated in any way to BeeLine Reader and its manufacturers.

https://doi.org/10.1016/j.chbr.2022.100197
Received 19 August 2021; Received in revised form 8 March 2022; Accepted 14 April 2022
Available online 20 April 2022

principles of reading. In fact, if we ignore current developments, we are ignoring the very nature of 'next generation reading'.

To avoid this pitfall, current and future research on the readability of digital texts should address the following issues. First, controlled experiments need to be carried out to determine which design principles of digital reading applications streamline reading and improve text comprehension – or hamper reading, for example because the applications increase the cognitive load of readers or distract readers, thereby inducing shallow, mindless reading (cf. Benedetto et al., 2015). Second, it is important to identify the types of readers that will benefit from texts with an alternate layout, for example because simplified layouts may be particularly beneficial for beginning or struggling readers but perhaps less so for more advanced readers (cf. Koornneef et al., 2019; Schneps, 2015; Schneps et al., 2010; Schneps, Thomson, Chen, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013). In the present study, we address these issues for the digital tool BeeLine Reader; its efficacy with respect to reading speed and comprehension accuracy is examined in two computerized reading experiments with young, beginning readers.

### *1.1. Streamlining the eye movements of beginning readers with BeeLine Reader?*

To become proficient readers, children must learn how to move their eyes over a text efficiently, to rapidly translate the abstract visual information on a page into a meaningful mental representation. Although the mechanisms that underlie the eye movements of developing readers are still not fully understood (Kim et al., 2019), it is well-known that the reading patterns of beginning readers differ from the patterns of (proficient) adult readers: (1) children read more slowly; (2) children make shorter saccades and display lower word skipping rates; (3) children fixate and re-fixate words more frequently; (4) children's fixation durations are longer; and (5) children show higher regression probabilities (for reviews see e.g., Blythe, 2014; Blythe & Joseph, 2011; Schroeder et al., 2015). Beginning readers also behave differently from adult readers when they move their eyes from the end of one line to the beginning of the next line. Although these return-sweep saccades have not received the same amount of attention in the literature as other aspects of eye-movement behaviour (Parker et al., 2019), several findings are noteworthy. First, children's return sweeps more often require a corrective saccade (Netchine et al., 1983). Second, beginning readers do not show a reduction in gaze duration for line-final words, but adult readers do (Tiffin-Richards & Schroeder, 2018). Third, relative to adult readers children launch their return sweeps closer to the right margin and their landing positions are closer to the left margin – yet they still initiate corrective saccades more frequently (Parker et al., 2019). Together these results suggest that return sweeps become more accurate and more efficient through practice and experience.

An appealing hypothesis is that the eye movements of beginning and struggling readers can be optimized (e.g., approaching a more adult-like pattern) by altering the layout of a text. As a result, additional cognitive resources should become available for higher-order comprehension processes such as monitoring, integration, and inference generation (e.g., Koornneef et al., 2019; Schneps et al., 2010; Schneps, Thomson, Chen, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013). BeeLine Reader was designed with that aim in mind. The colour gradients of BeeLine Reader should 'pull' the eyes from one line to the next by ensuring that the end of one line and the beginning of the next are coloured similarly (BeeLine Reader, 2021; see Fig. 1). The developers claim on their website that by using this 'simple cognitive trick' reading will be 'easier, faster, and more enjoyable' because it 'facilitates visual tracking and enables the reader to focus on other aspects of reading such as decoding and comprehension'. However, some words of caution are warranted.

First, although several scholars discuss BeeLine Reader in their work (Rayner et al., 2016; Rodrigue, 2017; Schneps, 2015) there are no published peer-reviewed studies on how BeeLine Reader affects the readability of texts and the eye-movement behaviour of readers.

Second, BeeLine Reader is designed to minimize the demands on visual tracking and oculomotor control during reading. However, there is an ongoing debate on whether these factors constrain reading development because difficulties in eye-movement behaviour may also stem from children's immature linguistic skills (Blythe, 2014; Blythe & Joseph, 2011; Huestegge et al., 2009; Schroeder et al., 2015). For example, according to Schroeder et al. (2015) children and adults do not differ much in terms of their efficiency of oculomotor control and we can only understand the developmental changes in eye movements if we know how cognitive systems associated with word recognition (and other aspects of language processing) mature and impact eye-movement behaviour.

Third, the colour gradients of BeeLine Reader generate a somewhat disorganized text layout and may hamper reading, for several reasons. For example, in the default settings of BeeLine Reader the colours red, blue, and black are used on a light-coloured (white) background. Legibility research on digital texts suggests that it may be better to use a black font on a white background than a blue or red font on a white background (Humar et al., 2008, 2014). In addition, there is some evidence that the use of different font colours within a word slows down reading, induces more reading errors, and impedes comprehension (Pinna & Deiana, 2018). Furthermore, research on coloured overlays suggests that the colour red has a detrimental effect on reading (Williams et al., 1992) – although blue overlays may improve reading (Iovino et al., 1998; Williams et al., 1992) and, moreover, studies on coloured overlays are controversial (Henderson et al., 2013). Another reason to be sceptical about the efficacy of BeeLine Reader is that the sections of a text that are coloured similarly are not necessarily linked in a meaningful way (e.g., readers may erroneously assume that all red sections in a text are connected to each other). As a final point, readers do not move their eyes smoothly over a text. Instead, eye movements are somewhat erratic and consist of rapid saccade-fixation sequences. In this view, it is not obvious how smooth transitions from one colour to the next would guide eye movements during reading.

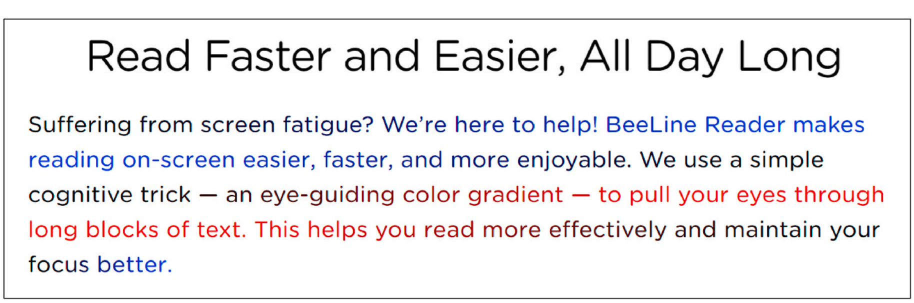


**Fig. 1.** An example of a BeeLine Reader layout as presented on the BeeLine Reader website (www.beelinereader.com). In addition to the default settings 'Bright' (i.e., gradients of black, blue, and red) as demonstrated here, readers can opt for other prearranged settings such as 'Blues' (black, blue, purple) and 'Grey' (solely gradients of grey) or customize the settings by choosing their own colour gradients. Reprinted from the BeeLine Reader website. Copyright 2017 by BeeLine Reader. Reprinted with permission. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

### 1.2. The present study

The main aim of the present study was straightforward: given the potential benefits and drawbacks of texts that are coloured in a gradient, we examined whether BeeLine Reader allows beginning readers to obtain a higher reading pace and improves their comprehension of a text, or hampers reading instead. In addition, we examined how the influence of BeeLine Reader on reading speed and comprehension interacts with the effects of other more traditional layout interventions that are used to improve the readability of texts for beginning readers, namely line spacing, line length, and text segmentation (we elaborate on these layout factors in Sections 2 and 3, respectively).

The reason for including additional layout manipulations was twofold. First, the effects of readability interventions are intertwined and many typographers stress that the efficacy of a layout variable cannot be assessed in isolation but should be studied together with other layout variables to avoid producing confounded results (Dyson, 2004; Joo et al., 2018). Second, by studying potential interaction effects, we can better contemplate how BeeLine Reader impacts reading. That is, if BeeLine Reader has a widespread positive impact on eye-movement behaviour, then you would expect that its benefits arise in many different text layouts. Alternatively, if BeeLine Reader specifically targets one aspect of eye-movement behaviour then you would expect more isolated effects – for example, if BeeLine Reader is particularly useful to make complex return sweeps more efficient, benefits should be observed more readily in dense layouts with long lines and little inter-line spacing (see Section 2 below).

Two experiments were conducted. In both experiments, we presented texts with a black font and texts with a BeeLine font (i.e., the letters were coloured in a gradient by adopting the algorithms of BeeLine Reader) to second- and third-grade readers in primary school. The reading times for the texts were recorded and each text was followed by a series of comprehension questions. In addition to this within-participants factor, both experiments included a between-participants factor. In Experiment 1, the effects of BeeLine Reader were examined for four layouts in which line length and inter-line spacing were varied (see Fig. 2). In Experiment 2, the effects of BeeLine Reader were examined for texts with a continuous layout (sentences continue on the same line as far as page width allows) and a segmented, discontinuous layout (line breaks in the middle of a sentence are removed as each sentence starts on a new line; see Fig. 6). Many scholars have pointed out that design decisions on the layout variables line spacing, line length, and segmentation will have an influence on the eye-movement behaviour of readers (e.g., Dyson, 2004; Evers-Vermeul, 2020; Koornneef et al., 2019; Levasseur et al., 2006; Schneps, Thomson, Sonnert, et al., 2013; Vanderschantz, 2008; Walker et al., 2018). Hence, examining the influence of BeeLine Reader on reading in combination with these layout variables will present interesting test cases for the hypothesis that the application 'facilitates visual tracking' and 'pulls the eyes from one line to the next' (BeeLine Reader, 2021).

## 2. Experiment 1

In Experiment 1, we examined the potential benefits and drawbacks of BeeLine Reader together with the layout variables line spacing and line length. A brief overview of several important principles and empirical findings that are related to these design features is provided below.

*Line spacing* is known to affect the readability of a text (Bernard et al., 2007; Blackmore-Wright et al., 2013; Calabrèse et al., 2010; Chung, 2004; Chung et al., 2008; Katzir et al., 2013; Ling & van Schaik, 2007; Madhavan et al., 2016; Rello et al., 2016). More specifically, reading pace increases as a function of enhanced line spacing, presumably by decreasing the adverse effect of visual crowding[2] between adjacent lines (Bernard et al., 2007; Chung, 2004). The advantages that are attributed to increased line spacing closely mirror the potential advantages of BeeLine Reader. First, it will be relatively easy for readers to keep their focus on the line that they are reading and, hence, they can avoid that their eyes move from a word on one line to the next word on adjacent lines (Vanderschantz, 2008). Second, it will also improve the accuracy of a return sweep as the target of this long-distance saccade is detected with less effort (Madhavan et al., 2016; Vanderschantz, 2008). Due to these advantages, low-vision patients, dyslectic readers, and children might benefit the most from increased line spacing (cf. Blackmore-Wright et al., 2013; Madhavan et al., 2016; Walker et al., 2018).

With respect to the typographic feature *line length* there is consensus that reading is disrupted if lines are either too long or too short. On the one hand, return sweeps are difficult if lines are too long because long lines increase the probability that readers will not arrive at the beginning of the next line, or end up on the wrong line (Dyson & Haselgrove, 2001; Nanavati & Bias, 2005; Rayner & Pollatsek, 1989). As a result, readers are forced to make a corrective eye movement to avoid re-reading the same line or omitting lines of text. Furthermore, long lines not only affect reading rate and eye movement patterns, but they potentially affect comprehension because complex return-sweep manoeuvres may have a detrimental effect on readers' concentration (Dyson & Haselgrove, 2001). On the other hand, inefficient reading may also occur if lines are too short. When reading short lines of text, readers are unable to establish regular rhythmic eye movements as they cannot exploit in full their peripheral vison and tend to decrease their saccade length, make more fixations, and increase the duration of these fixations (Dyson & Haselgrove, 2001; Rayner & Pollatsek, 1989). Based on these hypotheses on the advantages and drawbacks of long and short lines, moderate line lengths generally are recommended for optimal reading (Dyson & Haselgrove, 2001; Nanavati & Bias, 2005).

### 2.1. Goals and predictions

According to its developers, BeeLine Reader facilitates visual tracking and enables beginning readers to focus more on important aspects of reading such as decoding and comprehension. In a computerized reading experiment where second- and third-grade readers (6–9 years old) read texts with black and BeeLine fonts, we tested this hypothesis for four different layouts (see Fig. 2): (1) a layout with single-line spacing and a wide text window (abbreviated as Single-Wide); (2) a layout with single-line spacing and a narrow text window (Single-Narrow); (3) a layout with increased line spacing and a wide text window (Increased-Wide); (4) a layout with increased line spacing and a narrow text window (Increased-Narrow). We predicted that any beneficial effects of BeeLine Reader (i.e., faster reading, improved comprehension) should most likely occur for texts with a Single-Wide layout because this crowded configuration requires the most complex return-sweep saccades and increases the probability that readers lose track of the line that they are reading. No, or more modest, positive effects of BeeLine Reader were predicted for Single-Narrow and Increased-Wide layouts in which the settings for, respectively, line length and line spacing were better suited for beginning readers. Finally, BeeLine

[2] Visual crowding occurs when objects – notably abstract symbols such as letters – appear close together which makes it difficult to identify their individual features when perceived in the peripheral regions of the visual field (Pelli & Tillman, 2008). It is an important factor to take into account because crowding is intimately linked to the allocation of spatial attention (Petrov & Meleshkevich, 2011) and known to affect reading speed (Pelli & Tillman, 2008). This may hold in particular for children because there is evidence that they are influenced more by crowding phenomena than adults (Jeon et al., 2010).

Reader should be least effective for the layout that was fully customized for children, containing short lines and sufficient inter-line spacing (Increased-Narrow layout).

After completing the reading experiment, the children filled out a questionnaire in which they rated the use of black and gradient-coloured fonts on the following dimensions: comprehension, difficulty, convenience, and enjoyment. Thus, in addition to obtaining reading time and comprehension measures, we evaluated BeeLine Reader by probing readers' perceptions of the readability of gradient-coloured texts.

## 2.2. Materials and methods

### 2.2.1. Participants

Participants were 174 pupils (94 girls; mean age 7.4 years; range 6–9) in Grades 2 (98 children) and 3 from 9 primary schools in the Netherlands.[3] In both experiments reported in the present study, the children had no diagnosed behavioural and/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.

### 2.2.2. Design and procedure

All procedures were approved by the Leiden University Institute of Education and Child Studies ethics committee (project number ECPW-2019/237) and conducted in accordance with the Declaration of Helsinki. The freely accessible server (password protected) Ibex Farm (version 0.3.9; Drummond, 2018) and its supplementary software were used to run the reading experiment on a laptop or desktop at the schools of the participants. The experiment ran in the full-screen modus of an internet browser (Google Chrome) and consisted of two main blocks. Both blocks started with a practice text to familiarize the participants with the stimuli in each block. 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). The texts were presented in their entirety in a white (initially empty) text box on a blue background, using a sans-serif font. In one of the blocks, the texts were presented in a black font (Black condition). In the other block, the lines of the texts were coloured in a gradient by adopting the procedures of BeeLine Reader, ensuring that the end of one line and the beginning of the next were coloured similarly (BeeLine condition). In a BeeLine pathway the colour of the letters gradually changed from black to blue on the first line, from blue to black on the second line, from black to red on the third line, and from red to black on the fourth line – note that this cycle was repeated multiple times for each text (see Fig. 2 for examples of the BeeLine stimuli). During the experiment, the children were instructed to make a text appear on the computer screen by selecting a link labelled 'START' located directly below the text box. They did so by operating a mouse with their dominant hand, activating the link by clicking the left mouse button with their index finger. When the children finished reading the text, they clicked on the link again, which now was labelled as 'KLAAR' ('finished'), to proceed to the comprehension questions.[4] The elapsed time between the two mouse clicks was recorded to obtain the total reading time for a text. After each text, six comprehension questions appeared on screen, one by one. The test leader read out aloud the questions and recorded the answers of the child (see Fig. 3 for an overview of a single trial).

There were four versions of the experiment (see Fig. 2): (1) Single-Wide version with 42 participants; (2) Single-Narrow version with 44 participants; (3) Increased-Wide version with 44 participants; (4) Increased-Narrow version with 44 participants.[5] Line spacing was set to 1.0 lines for the single-line spacing conditions and to 1.5 lines for the conditions with increased line spacing. The texts with a narrow window covered 12–13 lines and each line contained on average 10.1 words. The texts with a wide window covered 7–8 lines and each line contained on average 15.9 words. In each version of the experiment, the ordering of the two experimental blocks and the four critical texts was rotated across four counterbalanced lists. Participants were randomly assigned to one of those lists. After completion of the reading task, the participants filled out a questionnaire (see Section 2.2.4 below). The questionnaire was presented digitally with Ibex Farm. The duration of a full test session was 20–40 min.

### 2.2.3. Texts and comprehension questions

The content of the texts was identical to the stimuli that were designed for a previous study (Koornneef et al., 2019). The stimuli consisted of six age-appropriate texts, including the two practice texts. The four critical texts consisted of two expository texts (one about the social structure of a community of lions and one about the human skeleton) and two 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 and their average length was 124 words (range: 117–132 words). To assess text comprehension, six 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 'gap-filling' inference; see e.g., Cain & Oakhill, 1999). The answers of the children were scored as correct or incorrect, based on a strict coding protocol.

### 2.2.4. Questionnaire

Using 5-point Likert scales, the participants provided separate ratings for the Black and BeeLine texts on four dimensions. First, they indicated how convenient it was for them to read Black and BeeLine texts (1 = not convenient at all; 5 = very convenient). Second, they rated how difficult it was for them to read Black and BeeLine texts (1 = not difficult at all; 5 = very difficult). Third, they rated how enjoyable it was for them to read Black and BeeLine texts (1 = not enjoyable at all; 5 = very enjoyable). Fourth, they rated their perceived level of comprehension for Black and BeeLine texts (1 = very low comprehension; 5 = very high comprehension).

## 2.3. Results

### 2.3.1. Reading experiment

For one participant, data of a trial in the Black condition (Single-Wide experiment) were incomplete. This trial (i.e., reading time for one text and answers to six comprehension questions) was removed from the dataset.

Table 1 and the left column of Fig. 4 report the results for the reading times (average reading time per word in milliseconds) and the accuracy scores for the comprehension questions (probability correct) in Experiment 1. Bayesian mixed-effects linear regression models were fitted to analyse the reading time data. The reading times were log transformed

[3] The initial sample consisted of 177 participants. The results of three participants were removed from the dataset because they did not follow the instructions of the test leaders.

[4] Before the experiment started, the test leader verified whether the participant was able to operate the mouse skilfully. If that was not the case, the test leader operated the mouse and participants indicated verbally whether they were ready to start reading the text and when they finished reading the text.

[5] The procedure for creating the stimuli that appeared in the text box was as follows. First, the different layout versions were created with Microsoft Word. Subsequently, the Word files were exported as PDF files and opened with the native Google Chrome BeeLine extension. For the BeeLine versions of the texts, we opted for the black-blue-red settings. To create the Black versions of the texts we customized the settings with all three colours set to black. High-resolution screenshots were taken to obtain the final stimuli that were presented with Ibex Farm.

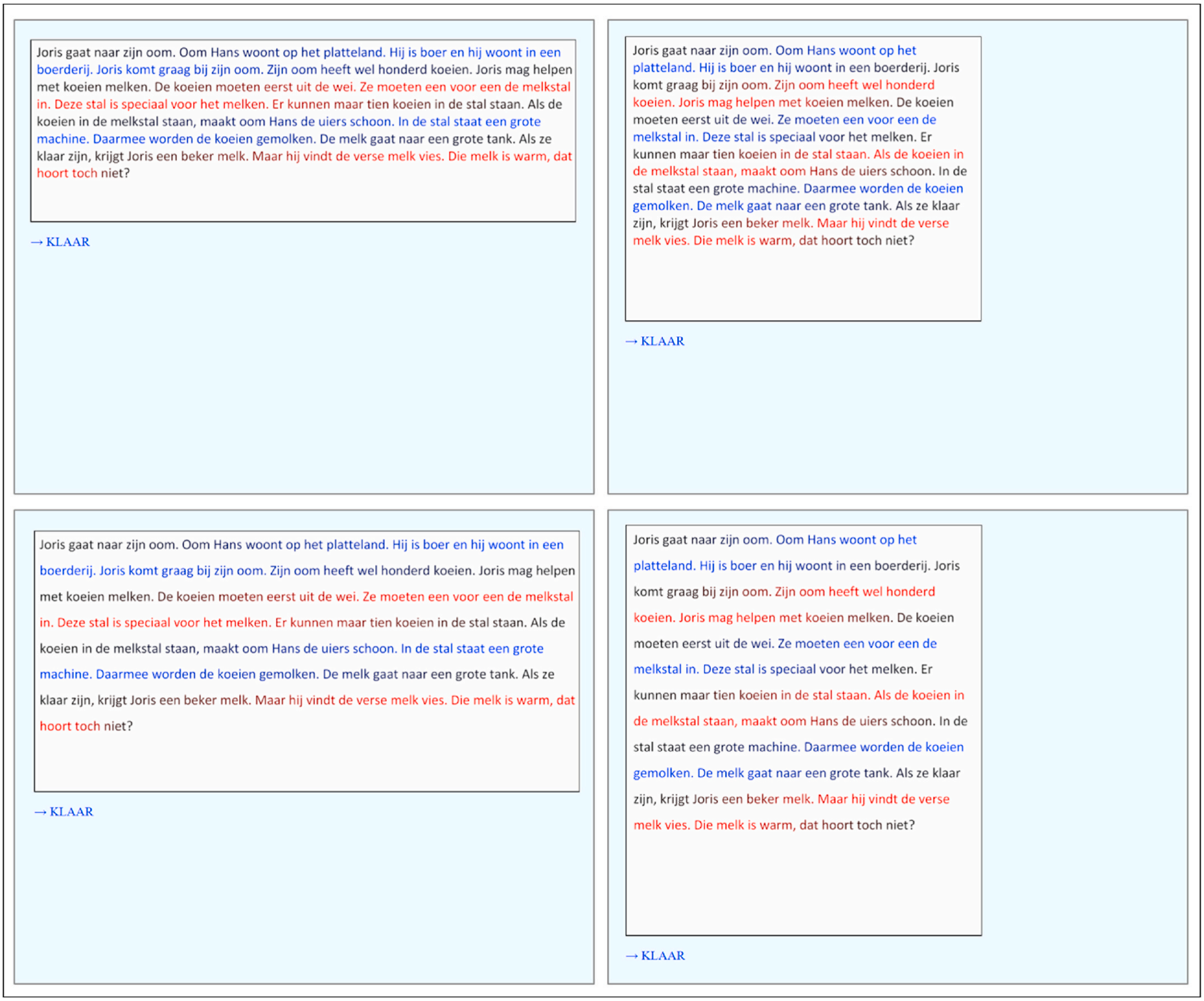


**Fig. 2.** Examples of BeeLine stimuli in Experiment 1. *Top left*. Layout with single-line spacing and a wide text window (Single-Wide condition). *Top right*. Layout with single-line spacing and a narrow text window (Single-Narrow condition). *Bottom left*. Layout with increased line spacing (i.e., 1.5 lines) and a wide text window (Increased-Wide condition). *Bottom right*. Layout with increased line spacing and a narrow text window (Increased-Narrow condition). A 2 × 4 mixed factorial design was constructed. Each participant read both Black and BeeLine texts (within-participants factor) but was assigned to only one layout version of the experiment (between-participants factor). The link 'KLAAR' ('finished') located below the text box was clicked by the participants to proceed to the comprehension questions.

to correct for right skewness. Bayesian mixed-effects logistic regression models were fitted to analyse the accuracy data (i.e., comprehension questions). The models were fitted with the statistical software R (Version 3.5.3) using the package Rstanarm (version 2.18.2). Participants and items were included as crossed random effects. Effect (i.e., deviation) coding was applied for the categorical independent variables. For the factor Font, the BeeLine condition was coded as 0.5 and the Black condition was coded as −0.5. For the factor Grade, Grade 3 was coded as 0.5 and Grade 2 was coded as −0.5. The factor Layout consisted of four levels (Single-Wide, Single-Narrow, Increased-Wide, Increased-Narrow) and a full contrast scheme consisted of three deviation contrasts. The Single-Wide version of the experiment was the baseline condition and was coded as −0.25 in each contrast. The coding of the other conditions depended on whether they were the target condition in a specific contrast. The target condition was coded as 0.75, and the remaining conditions were coded as −0.25. The models were computed in four chains consisting of 3000 iterations each. The weakly informative priors for the linear regression analyses (i.e., the analyses for the log-transformed reading times) were *Normal* ($\mu = 6.5$, $\sigma = 1.5$) for the intercept, and *Normal* ($\mu = 0$, $\sigma = 0.25$) for the coefficients. The weakly informative priors for the logistic regression analyses (i.e., the analyses for the comprehension questions) were *Normal* ($\mu = 0$, $\sigma = 3$) for the intercept, and *Normal* ($\mu = 0$, $\sigma = 1$) for the coefficients. Summaries (i.e., the median and the standard deviation of the median absolute difference) of all fixed effects in the models are provided in Table 2. As we optimized our design to study the influence of BeeLine and its interactions with other independent variables, we will only discuss the parameters that include the factor Font (i.e., the fixed effects of interest) and display their 50% and 95% credible intervals (percentile interval type; see the right column of Fig. 4). As a rule of thumb, we interpret the evidence for an effect as strong if zero lies outside the 95% credible interval (Kruschke et al., 2012; Nicenboim & Vasishth, 2016). For the parameters that revealed a main or modulating effect of Font, we will report the posterior probability mass that lies below or above zero (indicated as $P(\widehat{\beta} < 0)$ or $P(\widehat{\beta} > 0)$; see Nicenboim & Vasishth, 2016). Note that if zero is included within the 95% interval, there might still be

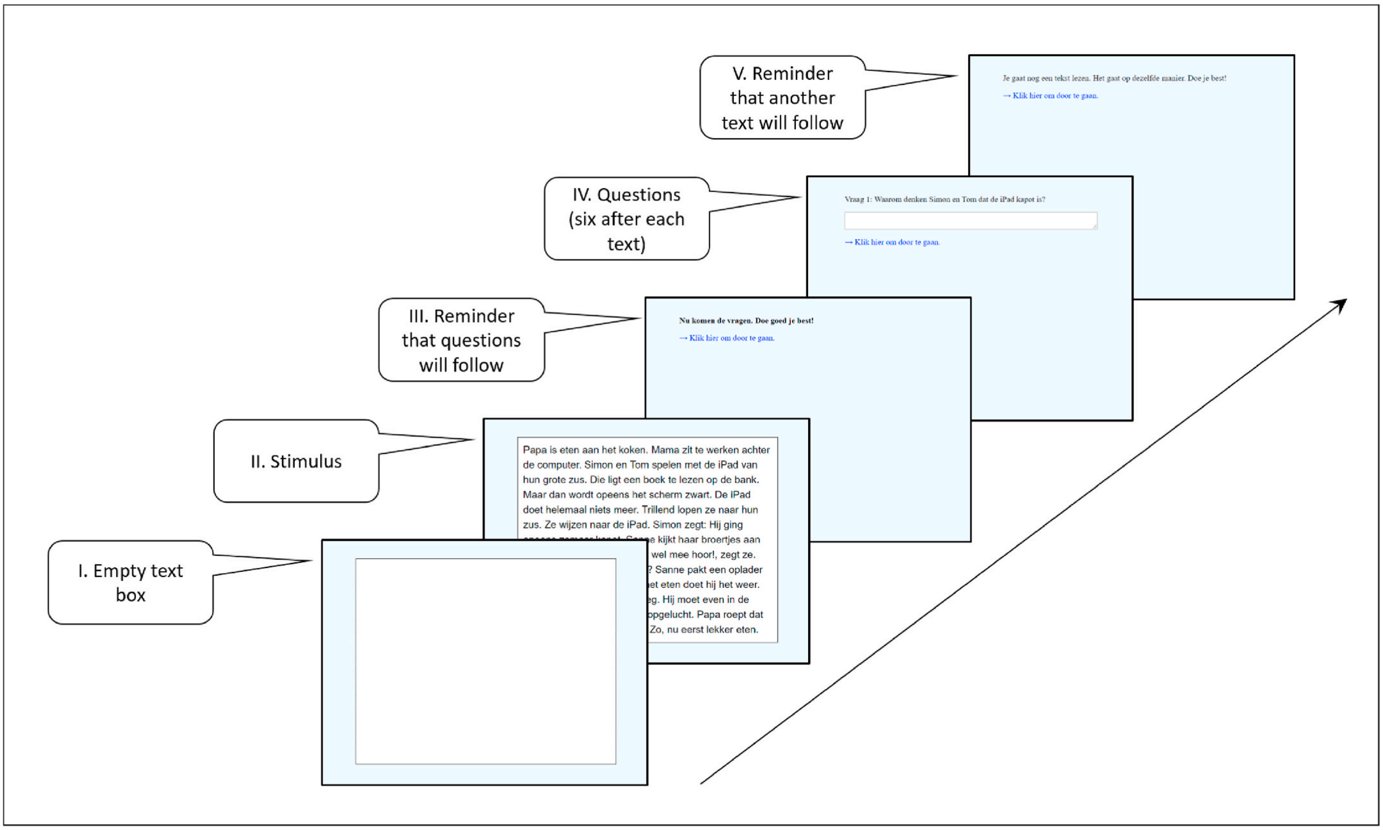


**Fig. 3.** Schematic overview of the time course of a trial. A trial started with an empty text box (I). When ready, participants clicked on a link below the text box with the label 'START' and the text appeared (II). When participants were finished reading the text, they clicked on a link with the label 'KLAAR' ('finished'). After a reminder (III), six comprehension questions appeared one by one on the screen (IV). A trial ended with a reminder that another text (i.e., a trial of the same type) was about to follow (V) – or new instructions were provided when the participants entered the second block of the experiment.

**Table 1**
Mean reading times for the texts (in milliseconds per word) and mean accuracy scores (probability correct) for the comprehension questions in Experiments 1 and 2. Standard errors are provided in parentheses.

| Experiment | Layout | Font | Grade 2 | | | | Grade 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Reading Time | | Accuracy | | Reading Time | | Accuracy | |
| 1 | Single-Wide | Black | 735 | (61) | .42 | (.033) | 493 | (18) | .70 | (.028) |
| | | BeeLine | 641 | (35) | .38 | (.032) | 480 | (22) | .71 | (.027) |
| | Single-Narrow | Black | 718 | (42) | .48 | (.028) | 493 | (21) | .68 | (.033) |
| | | BeeLine | 728 | (46) | .47 | (.028) | 498 | (24) | .59 | (.034) |
| | Increased-Wide | Black | 770 | (59) | .54 | (.029) | 566 | (46) | .57 | (.032) |
| | | BeeLine | 796 | (69) | .51 | (.029) | 582 | (49) | .56 | (.032) |
| | Increased-Narrow | Black | 599 | (31) | .45 | (.027) | 445 | (24) | .75 | (.031) |
| | | BeeLine | 622 | (31) | .48 | (.027) | 439 | (26) | .73 | (.032) |
| 2 | Discontinuous | Black | 708 | (39) | .62 | (.031) | 524 | (28) | .71 | (.025) |
| | | BeeLine | 656 | (43) | .58 | (.032) | 519 | (29) | .61 | (.027) |
| | Continuous | Black | 746 | (56) | .58 | (.032) | 547 | (27) | .56 | (.026) |
| | | BeeLine | 692 | (45) | .56 | (.032) | 531 | (22) | .61 | (.026) |

weak evidence for an effect, if the probability of the estimate being less than (or greater than) 0 is relatively large (Nicenboim & Vasishth, 2016). Hence, in some specific cases we tentatively interpret effects for which zero lies just inside the 95% credible interval.

The analyses for the reading time data showed a two-way interaction of Font x Layout (Font x Layout[SN]: $P(\widehat{\beta} > 0) = 0.99$; Font x Layout[IW]: $P(\widehat{\beta} > 0) > 0.99$; Font x Layout[IN]: $P(\widehat{\beta} > 0) > 0.99$) (see Table 2 and Fig. 4b). Pair-wise follow-up comparisons (see Fig. 4c) showed an effect of Font in the Single-Wide variant of the experiment (Median = −0.072, MAD-SD = 0.031, $P(\widehat{\beta} < 0) = 0.99$), indicating that the BeeLine condition induced faster reading than the Black condition. No credible effects of Font were observed in the other variants of the experiment (Single-Narrow: Median = 0.005, MAD-SD = 0.021, $P(\widehat{\beta} > 0) = 0.60$; Increased-Wide: Median = 0.019, MAD-SD = 0.024, $P(\widehat{\beta} > 0) = 0.80$; Increased-Narrow: Median = 0.017, MAD-SD = 0.022, $P(\widehat{\beta} > 0) = 0.78$). In addition, we observed a weak three-way interaction of Font x Layout x Grade when the Increased-Narrow layout was the target condition ($P(\widehat{\beta} < 0) = 0.97$). For second-grade pupils, pair-wise follow-up analyses (see Fig. 4d) showed that BeeLine texts induced faster reading than Black texts in the case of a Single-Wide layout (Median = −0.095, MAD-SD = 0.051, $P(\widehat{\beta} < 0) = 0.97$), yet BeeLine texts induced slower reading than Black texts in the case of an Increased-Narrow layout (Median = 0.055, MAD-SD = 0.029, $P(\widehat{\beta} > 0) = 0.97$). For third-grade pupils no clear

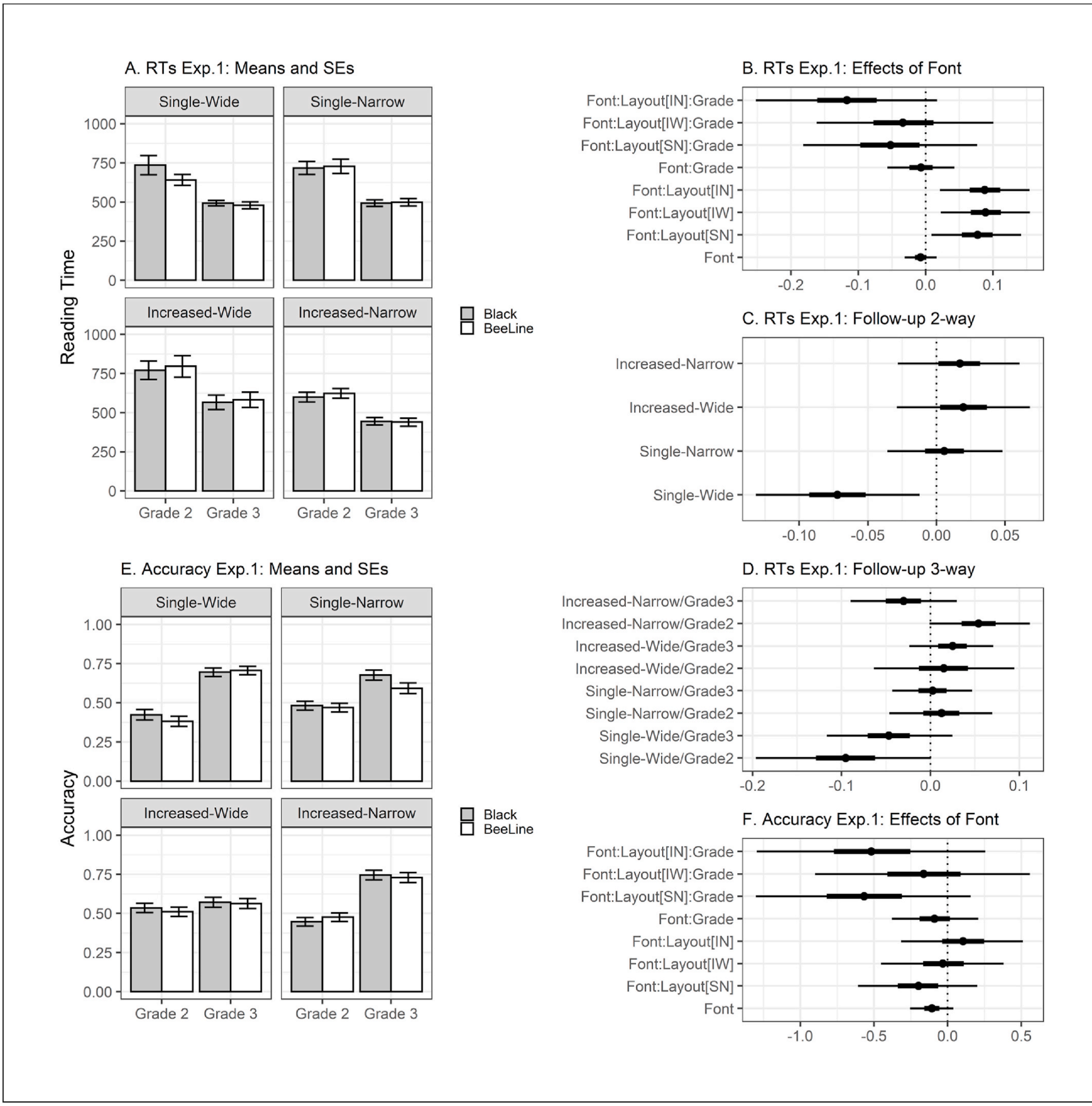


**Fig. 4.** Results of Experiment 1. In the left column, descriptive graphs of the data. In the right column, credible intervals (50% and 95% percentile intervals) for the estimates of the fixed effects that included the factor Font. Fig. A. Mean reading times and standard errors as a function of Font, Layout, and Grade. Fig. B. Credible intervals for the Bayesian mixed-effects analysis of the log-transformed reading times. Fig. C. Credible intervals for the pair-wise comparisons (Black vs. BeeLine) of the Font × Layout interaction that was observed in the reading time analysis. Fig. D. Credible intervals for the pair-wise comparisons (Black vs. BeeLine) of the Font x Layout × Grade interaction that was observed in the reading time analysis. Fig. E. Mean accuracy scores and standard errors as a function of Font, Layout, and Grade. Fig. F. Credible intervals for the Bayesian mixed-effects analysis of the accuracy scores. (RT = reading time per word; SE = standard error; [SN] = target condition is Single-Narrow layout; [IW] = target condition is Increased-Wide layout; [IN] = target condition is Increased-Narrow layout).

effects of Font emerged in any of the follow-up comparisons.

The analyses for the accuracy data showed no credible effects for the parameters of interest (see Table 2 and Fig. 4f).

### 2.3.2. Questionnaire

We will discuss the main effects of the factor Font.[6] Separate Bayesian mixed-effects linear regression models were fitted for the dimensions convenience, difficulty, enjoyment, and comprehension. The BeeLine condition was coded as 0.5 and the Black condition was coded as −0.5. Participants were included as a random effect. The models were computed in four chains of 3000 iterations with the following priors: *Normal* ($\mu = 3$, $\sigma = 1$) for the intercept and *Normal* ($\mu = 0$, $\sigma = 2$) for the coefficients. As showed in Fig. 5, BeeLine texts were rated as less convenient (Median = −0.471, MAD-SD = 0.127, $P(\widehat{\beta} < 0) > 0.99$) and more difficult (Median = 0.562, MAD-SD = 0.111, $P(\widehat{\beta} > 0) > 0.99$) to read than Black texts. The results for the remaining categories were less pronounced but pointed in a similar direction: BeeLine texts were less enjoyable to read (Median = −0.163, MAD-SD = 0.131, $P(\widehat{\beta} < 0) = 0.90$) and induced a lower level of perceived comprehension (Median = −0.182, MAD-SD = 0.112, $P(\widehat{\beta} < 0) = 0.95$) than Black texts did.

## 2.4. Summary of results

The influence of BeeLine Reader interacted with the combined influence of the layout features line spacing and line length. More specifically, BeeLine Reader allowed children to obtain a higher reading pace for a text format with single-line spacing and a wide text window – without impeding comprehension. This positive effect of BeeLine Reader was not observed for layouts with narrow text windows and/or sufficient line spacing. In fact, the analyses suggested that BeeLine Reader hampers second-grade pupils when they were reading texts in which both layout features were optimized for beginning readers (i.e., texts containing shorter lines *and* sufficient inter-line spacing). Hence, with that configuration the drawbacks of BeeLine Reader seem to outweigh its benefits. The analyses also suggested that second-grade readers were more sensitive to the influence of BeeLine Reader than third-grade readers. Other important findings of Experiment 1 were that BeeLine Reader did not affect text comprehension and that beginning readers preferred a black font over a gradient-coloured font – they rated BeeLine texts as more difficult and less convenient to read. We discuss these findings further in the General Discussion (Section 4).

# 3. Experiment 2

In Experiment 2, we assessed the potential benefits and drawbacks of BeeLine Reader in combination with a different layout variable: text segmentation. It is a popular design feature among publishers of schoolbooks who frequently opt for a layout in which texts are presented with each sentence starting on a new line of the page (cf. Evers-Vermeul, 2020; Land, 2009; van Silfhout, Evers-Vermeul, Mak, et al., 2014; van Silfhout, Evers-Vermeul, & Sanders, 2014). One of the advantages of these so-called 'discontinuous' texts is that the line and sentence endings coincide and together present a prominent cue to the reader that sentence wrap-up and integration processes should be initiated (Koornneef et al., 2019). In addition to stimulating sentence wrap-up, publishers assume that this layout is easier to read than a traditional continuous layout because it avoids that clausal units are interrupted by a line break, thereby limiting parsing problems for beginning readers during a return sweep (cf. Levasseur et al., 2006; Raban, 1982).

Two studies have examined this hypothesis for beginning readers, revealing mixed results. Whereas Evers-Vermeul (2020) reported that discontinuous texts are not felicitous for 7-year-old readers, Koornneef et al. (2019) showed that high-performing beginning readers perform better with a discontinuous layout than with a continuous layout. Interestingly, this does not hold for struggling beginning readers. They appear to perform better with a continuous layout than with a discontinuous layout. Koornneef et al. (2019) speculated that these differences between proficient and struggling beginning readers emerged because discontinuous texts pose higher demands on eye-movement control processes than continuous texts do – in contrast to the claims made by some publishers. For example, if the width of the text window is held constant across conditions, discontinuous texts will cover more lines than the same texts presented in a continuous fashion. As a result, more return-sweep saccades are required, which increases the demands on oculomotor control processes (Evers-Vermeul, 2020; Koornneef et al., 2019). Furthermore, two other issues (which have gone unnoticed in the literature cited above) strengthen the idea that discontinuous texts are perhaps more difficult than continuous texts. First, return-sweep saccades are more erratic in discontinuous layouts than in continuous layouts because a discontinuous layout induces poorly ragged margins, creating distracting shapes of white space on the right side of the page (see Fig. 6). Second, return-sweep saccades that end up on the wrong line of a discontinuous text may not always be identified immediately as erroneous because they do not cause a parsing problem for the reader – note that inaccurate return sweeps in a continuous layout often will result in syntactic processing difficulties that signal to the reader that something went wrong during reading.[7]

Hence, although there are several open issues in the debate on the advantages and disadvantages of continuous and discontinuous layouts, scholars agree on the idea that these layouts pose different demands on eye-movement behaviour, most notably the processes that control how readers move their eyes from one line of text to the next. In that view, continuous and discontinuous layouts present salient test cases for expanding our research on the efficacy of BeeLine Reader.

## 3.1. Goals and predictions

Experiment 2 was conducted to replicate and extend the findings of Experiment 1. Again, we presented texts with black and BeeLine fonts to second- and third-grade readers as a within-participants factor. The layout variable text segmentation was included as a between-participants factor; in one version of the experiment the texts were formatted in a discontinuous layout and in the other version the texts were formatted in a continuous layout (see Fig. 6). Based on our findings in Experiment 1, we made the following predictions. First, it was predicted that if BeeLine Reader improved the reading speed of beginning readers, these effects should emerge more readily for second-grade readers than for third-grade readers. Second, no comprehension advantage was predicted for BeeLine Reader (neither for second-grade readers nor for third-grade readers). Furthermore, we speculated that return-sweep manoeuvres are more demanding in a discontinuous layout than in a continuous layout. Based on this idea, a third prediction was that beneficial effects of BeeLine Reader should occur more prominently in a discontinuous layout than in a continuous layout.

## 3.2. Materials and methods

### 3.2.1. Participants

Participants were 97 pupils (52 girls; mean age 8.1 years; range: 6–9) in Grade 2 (40 children) and Grade 3 from 8 primary schools in the Netherlands. None of them participated in Experiment 1.

[6] We also fitted more complex models that included the factors Layout and Grade. These models revealed no reliable interactions.

[7] We thank Jacqueline Evers-Vermeul for pointing out this potential disadvantage of discontinuous texts.

**Table 2**
Summaries of the Font x Layout x Grade Bayesian mixed-effects models in Experiment 1. On the left, fixed-effect estimates and the associated statistics for the log-transformed reading times. On the right, fixed-effect estimates (on a logit scale) and the associated statistics for the accuracy scores. These models are discussed in Section 2.3.1 and further illustrated in Fig. 4. (SD = standard deviation; MAD-SD = standard deviation of the median absolute difference; [SN] = target condition is Single-Narrow layout; [IW] = target condition is Increased-Wide layout; [IN] = target condition is Increased-Narrow layout).

| Fixed Effects | Reading Time | | Accuracy | |
|---|---|---|---|---|
| | Median | MAD-SD | Median | MAD-SD |
| (Intercept) | 6.302 | 0.038 | 0.373 | 0.226 |
| Font | −0.007 | 0.013 | −0.107 | 0.075 |
| Layout[SN] | 0.029 | 0.077 | −0.016 | 0.231 |
| Layout[IW] | 0.088 | 0.077 | −0.085 | 0.227 |
| Layout[IN] | −0.092 | 0.081 | 0.229 | 0.236 |
| Grade | −0.284 | 0.060 | 1.047 | 0.174 |
| Font:Layout[SN] | 0.076 | 0.035 | −0.198 | 0.208 |
| Font:Layout[IW] | 0.089 | 0.035 | −0.032 | 0.204 |
| Font:Layout[IN] | 0.088 | 0.036 | 0.107 | 0.216 |
| Font:Grade | −0.007 | 0.024 | −0.089 | 0.148 |
| Layout[SN]:Grade | 0.004 | 0.132 | −0.526 | 0.422 |
| Layout[IW]:Grade | 0.018 | 0.128 | −1.092 | 0.415 |
| Layout[IN]:Grade | 0.015 | 0.132 | 0.099 | 0.439 |
| Font:Layout[SN]:Grade | −0.052 | 0.066 | −0.560 | 0.388 |
| Font:Layout[IW]:Grade | −0.033 | 0.067 | −0.153 | 0.371 |
| Font:Layout[IN]:Grade | −0.117 | 0.067 | −0.507 | 0.406 |
| Random Effects | | | | |
| Groups | Name | SD | Name | SD |
| Participants | (Intercept) | 0.386 | (Intercept) | 1.002 |
| Items | (Intercept) | 0.062 | (Intercept) | 1.071 |
| Nr. Observations | 695 | | 4170 | |

### 3.2.2. Texts and comprehension questions

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

### 3.2.3. 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. The main set-up and procedures were identical to Experiment 1. Each participant read three texts with a Black font (one practice text and two test texts), three texts with a BeeLine font (one practice text and two test texts), and we applied the same procedure for a single trial (see Fig. 3 for an overview). There were two versions of Experiment 2. In one version (47 participants), line breaks in the middle of a sentence were removed as each sentence started on a new line (Discontinuous layout, see Fig. 6, left side). In the other version (50 participants), sentences continued on the same line as far as page width allowed (Continuous layout, see Fig. 6, right side).[8] The texts with a discontinuous layout consisted of 19 lines and each line contained on average 6.5 words. The texts with a continuous layout consisted of 11–12 lines and each line contained on average 10.5 words. In both versions of the experiment, the ordering of the two experimental blocks and the four critical texts was rotated across four counterbalanced lists. Participants were randomly assigned to one of those lists. The duration of a full test session was 20–35 min.

[8] The textual stimuli that appeared in the text box were created by copy-and-pasting the content of the texts into the 'BeeLine pasteboard' (http://www.beelinereader.com/pasteboard.html). We took high-resolution screenshots of the pasteboard to obtain the final image files that were presented with Ibex Farm.

## 3.3. Results

Table 1 and the left column of Fig. 7 report the results for the reading times and accuracy scores of the comprehension questions in Experiment 2. Bayesian mixed-effects models with the same priors as in Experiment 1 were computed to analyse the data. Participants and items were included as crossed random effects. Deviation coding was used for the categorical independent variables. For the factor Font, the BeeLine condition was coded as 0.5 and the Black condition was coded as −0.5. For the factor Layout, the Continuous layout was coded as 0.5 and the Discontinuous layout was coded as −0.5. For the factor Grade, Grade 3 was coded as 0.5 and Grade 2 was coded as −0.5. The models were computed in four chains consisting of 2000 iterations each.

The analyses for the reading time data showed an effect of Font ($P(\widehat{\beta} < 0) > 0.99$) and a two-way interaction of Font x Grade ($P(\widehat{\beta} > 0) = 0.98$) (see Table 3 and Fig. 7b). The main effect of Font indicates that the BeeLine conditions induced faster reading than the Black conditions. Follow-up analyses of the two-way interaction (see Fig. 7c) revealed that this effect of Font was present for second-grade readers (Median = −0.075, MAD-SD = 0.023, $P(\widehat{\beta} < 0) > 0.99$), yet no credible effect was observed for third-grade readers (Median = −0.018, MAD-SD = 0.017, $P(\widehat{\beta} < 0) = 0.86$).

The analyses for the accuracy data showed an effect of Font ($P(\widehat{\beta} < 0) = 0.95$), a two-way interaction of Font x Layout ($P(\widehat{\beta} > 0) > 0.99$), and a (weak) three-way interaction of Font x Layout x Grade ($P(\widehat{\beta} > 0) = 0.96$) (see Table 3 and Fig. 7e). The main effect indicates that the BeeLine conditions induced lower accuracy scores than the Black conditions did. Follow-up analyses of the interactions (see Fig. 7f) showed that this effect of Font was carried by the scores of third-grade readers in the Discontinuous condition (Median = −0.73, MAD-SD = 0.21, $P(\widehat{\beta} < 0) > 0.99$) because no credible effects of Font were observed in the other comparisons of interest (Discontinuous/Grade 2: Median = −0.21, MAD-SD = 0.23, $P(\widehat{\beta} < 0) = 0.82$; Continuous/Grade 2: Median = −0.001, MAD-SD = 0.23, $P(\widehat{\beta} < 0) = 0.50$; Continuous/Grade 3: Median = 0.19, MAD-SD = 0.18, $P(\widehat{\beta} > 0) = 0.85$).

## 3.4. Summary of results

The reading time analyses showed that BeeLine Reader texts increased the reading speed of second-grade readers but did not affect the reading speed of third-grade readers. Furthermore, BeeLine Reader did not improve comprehension. In fact, relative to their control texts, comprehension of BeeLine texts with a discontinuous layout was *impeded* for third-grade readers. The prediction that beneficial effects of BeeLine Reader should occur more prominently in a discontinuous layout than in a continuous layout was not confirmed by the data. The implications of these findings will be discussed in the General Discussion below.

# 4. General Discussion

In two reading experiments with primary school pupils, we examined the claims of the developers of BeeLine Reader that smooth colour transitions in texts increase reading speed, improve comprehension, and make reading more enjoyable. To obtain a comprehensive picture we tested these hypotheses for second- and third-grade pupils, assessing the influence of BeeLine Reader in several layouts in which we varied features that are known to impact the reading processes of beginning readers (line spacing, line length, text segmentation) (see e.g., Evers-Vermeul, 2020; Koornneef et al., 2019; Nanavati & Bias, 2005; Vanderschantz, 2008; Walker et al., 2018). The experiments revealed that BeeLine Reader can increase the reading speed of beginning readers. However, this positive influence depends on characteristics of the reader and the text. Reading time advantages for BeeLine texts primarily

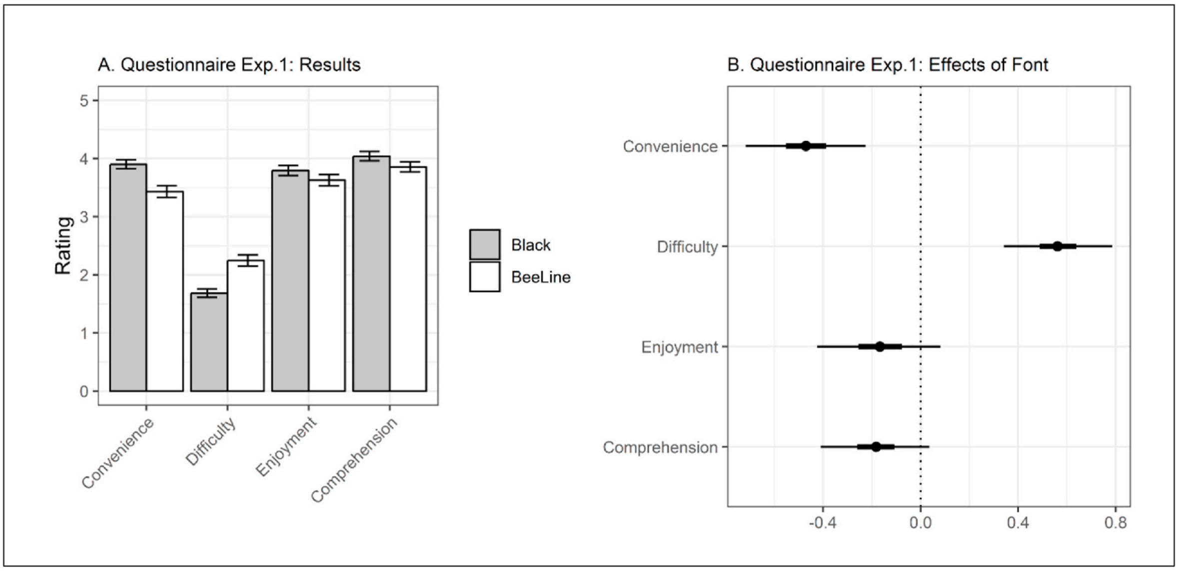


**Fig. 5.** Results of the questionnaire in Experiment 1. Fig. A. Mean ratings and standard errors for the four dimensions of the questionnaire as a function of FONT (convenience: 1 = not convenient at all, 5 = very convenient; difficulty: 1 = not difficult at all, 5 = very difficult; enjoyment: 1 = not enjoyable at all, 5 = very enjoyable; comprehension*:* 1 = very low comprehension, 5 = very high comprehension). Fig. B. Credible intervals for the Bayesian mixed-effects analyses of the factor FONT on the four dimensions.


**Fig. 6.** Examples of BeeLine stimuli in Experiment 2. On the left, a Discontinuous text. On the right, a Continuous text. A 2 × 2 mixed factorial design was constructed. Each participant read both Black and BeeLine texts. There were two versions of the experiment: a participant read either Discontinuous texts or Continuous texts. We ensured that all the words in the BeeLine-Discontinuous and BeeLine-Continuous versions of a text were coloured in the same way. As a result, the BeeLine-Discontinuous versions did not fully adhere to the default algorithms of BeeLine Reader. For example, a default BeeLine cycle spans four lines of text (see Continuous text on the right) whereas our customized BeeLine cycle for the Discontinuous versions spans more lines of text (e.g., about six lines in the text on the left).

emerge for pupils in second grade (not in third grade) when they read texts with a high visual density ('crowded' texts consisting of wide windows with little inter-line spacing). When second-grade readers process texts that are optimized for their reading level (texts with narrow windows and increased inter-line spacing) they tend to display a slower reading pace in texts with a BeeLine font. Concerning the developers' second claim: BeeLine Reader does not improve the mental representation of a text and it can even have a negative impact on comprehension. Our data are also incompatible with the third hypothesis of the developers as the participants do not rate BeeLine texts as more enjoyable to read than texts with a black font. In fact, the children indicate that BeeLine texts are more difficult and less convenient to process. We elaborate on each of these findings below.

### 4.1. BeeLine Reader and reading speed

The claim made for BeeLine Reader is that it may facilitate visual tracking, thereby streamlining and accelerating the eye movements of readers. The developers of BeeLine Reader do not explicitly discuss the goals of their application in terms of first-pass fixations, re-reading fixations, saccades, regressions, and return sweeps, but they do seem to suggest that the application may have a positive impact on several, if not most, aspects of eye-movement behaviour. This may hold for the straightforward reason that different eye-movement measures index correlated and inherently intertwined processes. For example, when return-sweep saccades become more accurate this will also affect more global reading patterns as fewer (regressive) eye movements are required to repair erroneous transitions from one line to the next. This could decrease re-reading durations and, moreover, first-pass processing may be optimized (i.e., longer saccades, shorter fixation durations) because readers get into a more comfortable and swift reading mode.

However, in our opinion it is not likely that BeeLine Reader has a positive impact on all these aspects of eye-movement behaviour. A notable disadvantage of a gradient-coloured font is that the letters within a word are coloured differently which may have a negative impact on decoding processes of individual words, with slower reading as a result (cf. Pinna & Deiana, 2018). In addition, it is not obvious how smooth transitions from one colour to the next would guide saccadic eye movements which do not follow a smooth, continuous path from left to right on a line of text. Taking these limitations of the application into account, BeeLine Reader most likely will not support *intra*-linear eye movements during reading. Instead, a more plausible explanation of the data is that BeeLine Reader can have a positive influence on *inter*-linear eye movements. The return sweep manoeuvres that readers generate to shift their view to the next line of a text will become easier due to the visual anchors offered by BeeLine Reader. As a result, BeeLine Reader is more effective for second-grade readers than for third-grade readers because they have less experience with texts consisting of multiple lines and encounter more difficulties in planning and executing return sweeps than third grade readers do. It also explains why benefits of BeeLine Reader arise for texts with long lines and little inter-line spacing but hampers processing of texts with short lines and sufficient spacing. Whereas in the former type of texts return sweeps are especially difficult, this is not the case in the latter type of texts in which the layout features line length and line spacing are modified to simplify return sweeps for inexperienced readers (e.g., Nanavati & Bias, 2005; Vanderschantz, 2008; Walker et al., 2018).

This bird's-eye view on the reading time data seems to reveal a clear picture of the impact of Beeline Reader on the processing of texts. However, a more detailed inspection of the reading times within and between experiments reveals two puzzling findings. First, based on the idea that poorly ragged margins in a text increase return-sweep difficulty, we reasoned that the positive effects of BeeLine Reader should emerge more readily for discontinuous texts than for continuous texts. This hypothesis was not confirmed by the data because similar improvements of reading speed were observed across these layouts for second-grade readers (Experiment 2). There are several possible scenarios to account for the absence of the anticipated interaction effect. Assuming that BeeLine Reader streamlines return-sweep saccades, a first possibility is that we overestimate the influence of poorly ragged margins on the complexity of planning and executing return sweeps during reading. Alternatively, it is possible that we underestimate the challenges that continuous layouts pose to return-sweep processes. For example, it has been suggested that return sweeps in continuous texts are difficult because clausal units are often interrupted by a line break, which may induce parsing problems and slows down processing for beginning readers (cf. Levasseur et al., 2006; Raban, 1982). Furthermore, methodological considerations in our study may have reduced the efficacy of BeeLine Reader for texts with a discontinuous layout: the lines in a discontinuous layout were on average much shorter than the lines in a continuous layout (6.5 vs. 10.5 words per line) and to ensure that all the words in the BeeLine-discontinuous and the BeeLine-continuous versions of a text were coloured in the same way, the former versions did not fully adhere to the default algorithms of BeeLine Reader (see Fig. 6). In reflecting on these scenarios, we can merely state that our findings are inconclusive regarding the potential differential effects of BeeLine Reader in continuous and discontinuous layouts.

Second, readers displayed a reading time advantage for BeeLine texts with a continuous layout (Experiment 2), but such advantage was not observed for BeeLine texts with a narrow window and single-line spacing (Experiment 1). These layouts were almost identical (the texts with a continuous layout covered 11–12 lines and each line contained on average 10.5 words; the texts with a single-narrow layout covered 12–13 lines and each line contained on average 10.1 words). We do not have an explanation for the mixed findings across experiments although subtle variations in the stimuli and samples of participants could play a role here. These findings emphasize that an accelerating effect of BeeLine Reader on reading speed is not a robust but a highly context-dependent phenomenon.

### 4.2. BeeLine Reader and reading comprehension

An important purpose of reading is to construct a coherent mental representation of a text (e.g., Graesser et al., 1994; Kintsch, 1988; van den Broek, 1988). The developers of BeeLine Reader claim that the application improves the quality of such mental models. The reason for this hypothesis could be twofold. On the one hand, BeeLine Reader may have a *direct* positive influence on comprehension because the application decreases the probability that readers lose track of the lines they are reading – most notably after a return-sweep saccade. As a result, it should be easier to integrate the meaning of sentences into a unified, coherent mental representation. On the other hand, BeeLine Reader may also have an *indirect* positive influence on reading comprehension. Because BeeLine Reader is thought to optimize the eye movements of (beginning) readers, additional cognitive resources become available for higher-order comprehension processes such as monitoring, integration, and inference generation, thereby improving the quality of readers' mental models of texts (cf. Koornneef et al., 2019; Schneps et al., 2010; Schneps, Thomson, Chen, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013).

Our findings do not support these ideas. Although in some cases BeeLine texts induce more efficient reading (i.e., faster reading while preserving comprehension) none of the analyses revealed a comprehension advantage for BeeLine texts in comparison to texts with a black font. The only credible effect on reading comprehension was observed in Experiment 2, with *lower* comprehension scores for third-grade readers when they read discontinuous texts with a BeeLine font. We should be cautious with this single finding as it emerged as part of a subtle interaction effect (i.e., it could reflect a false-positive finding), but it does underscore the premise that adding gradient colours to a text may not be the best way to improve reading comprehension. Colours may distract readers – specifically the colour red may have a detrimental effect (cf. Humar et al., 2014, 2008; Pinna & Deiana, 2018; Williams et al., 1992) – or they may even provoke misinterpretations if readers erroneously assume that sections in a text with a similar colour are semantically and/or referentially connected to each other.

### 4.3. Beginning readers' opinions about BeeLine Reader

Next to the hypotheses that BeeLine Reader improves reading speed and reading comprehension, we examined whether the application makes reading easier and more enjoyable in the eyes of beginning readers. If that premise holds, BeeLine Reader could be used to increase motivation and feelings of self-efficacy for the act of reading. Such reader characteristics are precursors for becoming skilled readers later in life and it is therefore important to maximize these traits during reading acquisition (cf. Toste et al., 2020; Unrau et al., 2018). However, the data of the questionnaire (Experiment 1) show that beginning readers rate BeeLine texts as more difficult and less convenient to read

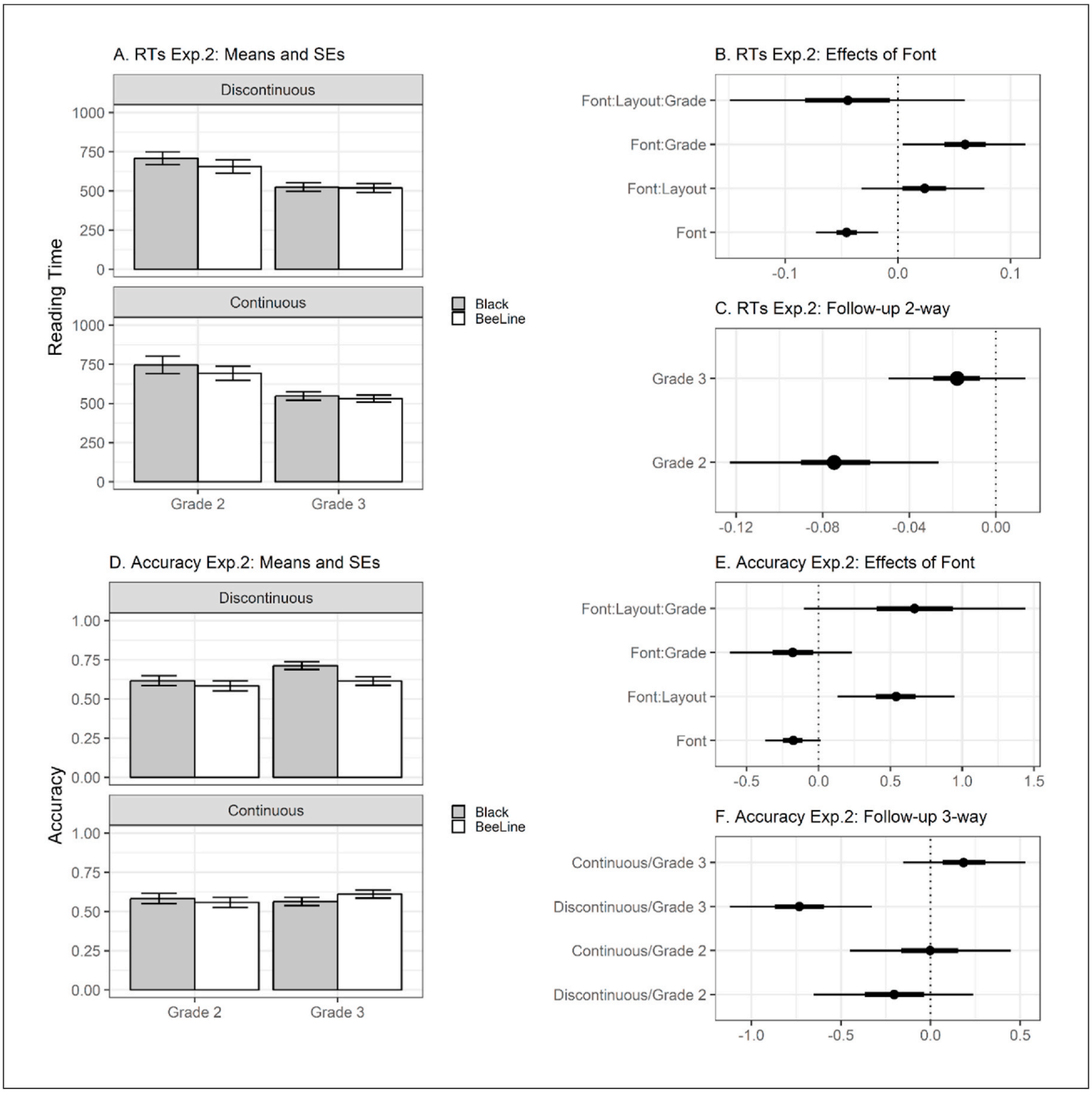


**Fig. 7.** Results of Experiment 2. In the left column, descriptive graphs of the data. In the right column, credible intervals (50% and 95% percentile intervals) for the estimates of the fixed effects that included the factor Font. Fig. A. Mean reading times and standard errors as a function of Font, Layout, and Grade. Fig. B. Credible intervals for the Bayesian mixed-effects analysis of the log-transformed reading times. Fig. C. Credible intervals for the pair-wise comparisons (Black vs. BeeLine) of the Font × Grade interaction that was observed in the reading time analysis. Fig. D. Mean accuracy scores and standard errors as a function of Font, Layout, and Grade. Fig. E. Credible intervals for the Bayesian mixed-effects analysis of the accuracy scores. Fig. F. Credible intervals for the pair-wise comparisons (Black vs. BeeLine) of the Font x Layout × Grade interaction that was observed in the accuracy score analysis. (RT = reading time per word; SE = standard error).

than texts with a black font. Similar (but weaker) trends are observed when the children are asked whether they enjoyed reading the texts and how well they understood the content of the texts: on these dimensions children also seem to prefer a black font over a gradient-coloured font. This robust, multi-dimensional preference for texts with a black font may reflect a familiarity effect. The children are used to reading in a black font and they may favour a text format that most closely resembles the formats of the books that they read in school and at home. Anecdotal evidence of the questionnaire seems to be in line with this idea. The questionnaire included an optional question where the children were invited to share their thoughts on the purpose of BeeLine Reader. Many children mentioned that gradient-colours were used to distract them during reading, and they indicated that they preferred a 'normal' text.

### 4.4. Limitations of the study and future directions

Digital reading applications should not be ignored in scientific research because they offer new and fruitful windows into applied and fundamental questions on reading acquisition. Our study contributes to this endeavour but many open issues remain. We highlight several key issues below.

An important premise in our research is that BeeLine Reader streamlines return-sweep manoeuvres during reading. Although our data are partly consistent with this idea, the methods that we use do not provide a direct test of the hypothesis. Eye tracking is a powerful technique to address this limitation. It offers high temporal resolution and allows researchers to extract numerous dependent measures to study the influence of BeeLine Reader on first-pass reading, re-reading, saccades, regressions, return sweeps, and so on. As a result, it provides a direct test of whether any advantages of BeeLine Reader should be attributed to

**Table 3**
Summaries of the Font x Layout x Grade Bayesian mixed-effects models in Experiment 2. On the left, fixed-effect estimates and the associated statistics for the log-transformed reading times. On the right, fixed-effect estimates (on a logit scale) and the associated statistics for the accuracy scores (i.e., performance on comprehension questions). These models are discussed in Section 3.3 and further illustrated in Fig. 7. (SD = standard deviation; MAD-SD = standard deviation of the median absolute difference).

| Fixed Effects | Reading Time | | Accuracy | |
|---|---|---|---|---|
| | Median | MAD-SD | Median | MAD-SD |
| (Intercept) | 6.342 | 0.045 | 0.653 | 0.304 |
| Font | −0.045 | 0.014 | −0.179 | 0.110 |
| Layout | 0.036 | 0.068 | −0.301 | 0.230 |
| Grade | −0.237 | 0.070 | 0.201 | 0.235 |
| Font:Layout | 0.024 | 0.027 | 0.528 | 0.211 |
| Font:Grade | 0.059 | 0.028 | −0.185 | 0.207 |
| Layout:Grade | 0.005 | 0.126 | −0.215 | 0.426 |
| Font:Layout:Grade | −0.042 | 0.055 | 0.677 | 0.392 |
| Random Effects | | | | |
| Groups | Name | SD | Name | SD |
| Participants | (Intercept) | 0.349 | (Intercept) | 1.017 |
| Items | (Intercept) | 0.062 | (Intercept) | 1.392 |
| Nr. Observations | 388 | | 2328 | |

enhanced *intra*-linear eye movements, to enhanced *inter*-linear eye movements, or to both.

Several other methodological aspects of our study should be considered. First, we neither explained the purpose of BeeLine Reader to the participants in advance nor did we include a training session to familiarize them with the application (although we did include a BeeLine practice text). We opted for this approach to avoid inducing biased opinions about the application. Second, the test sessions were kept short for the children to minimize any confounding effects due to fatigue or loss of concentration. A disadvantage of these methodological considerations is that we could be either overestimating or underestimating the beneficial effects of BeeLine Reader. On the one hand, the children may be attracted to the colours in the BeeLine texts which raises the question of whether beneficial effects emerge due the application itself or due to a novelty effect – which may wear off quickly. On the other hand, it is also possible that the full potential of BeeLine Reader becomes visible after providing more detailed instructions and offering training sessions to its users. Hence, future studies should consider using a (longitudinal) research design to track the benefits and drawbacks of BeeLine Reader over a longer period of time to resolve this ambiguity.

In the current study we examined the reading processes of typically developing beginning readers for short, 'static' (i.e., scrolling was prohibited) BeeLine texts. So, in addition to future directions as presented above, there are many other ways to extend the research on the effectiveness of BeeLine Reader. First, not only typically developing beginning readers may benefit from the features of BeeLine Reader but similar or different (dis)advantages may occur for other populations of readers, such as healthy adolescent and adult readers, people with a visual impairment, struggling (e.g., dyslectic) readers, older adults, and so on. Second, the influence of BeeLine Reader on longer and more complex texts (of various genres) should be examined. Third, because BeeLine Reader is developed to enhance the reading process of modern-day digital texts, future studies should include stimuli with hyperlinks and allow readers to scroll through texts. Fourth, in the current study we used the default settings of BeeLine Reader consisting of gradients of black, blue, and red. We postulated that especially the colour red may have a detrimental effect on reading. However, BeeLine Reader also offers settings without the colour red (black-blue-purple, only gradients of grey, fully customized settings) and these alternative settings may be more effective than the default settings.[9] Fifth, a more fundamental issue for future research is to explore whether and how the algorithms of BeeLine Reader affect the degree of visual crowding in a text. Sixth, BeeLine Reader is just one of many digital reading tools that deserve more and finer-grained attention from the scientific world (Spritz, Immersive Reader, Span Limiting Tactile Reinforcement, ReadMe!, WebClipRead, and so forth).

### 4.5. Concluding remarks

Based on the findings in the current study we conclude that BeeLine Reader might be useful, but only for *some* beginning readers in *some* situations. Potential users should keep in mind that BeeLine Reader is a charged application with no clear theoretical underpinnings and may have a detrimental effect on reading speed, reading comprehension, and more subjective readability measures. That said, the observation that BeeLine Reader induces more efficient reading for second-grade readers implies that it is not inconceivable that (in the future) BeeLine Reader can be a useful tool in educational settings; for example, as a scaffolding technique to support the early phases of reading acquisition, when beginning readers practice to maintain their focus on the correct line of text and familiarize themselves with the default directional rules of reading (left to right, top to bottom). As such, BeeLine Reader may reflect a contemporary substitute for traditional 'finger tracking' techniques (i.e., the movement of a child's index finger that points to printed text while reading) (e.g., Rodrigue, 2017). However, whether BeeLine Reader should be implemented to foster reading (acquisition) processes, is clearly an open issue. Moreover, decisions on implementing BeeLine Reader as an educational tool not only depend on learning gains in the short run, but also on learning gains in the long run. For example, presenting texts with BeeLine layouts to beginning readers may ultimately hamper their reading skills for texts with a traditional layout – just like learning how to ride a bicycle by using lateral training wheels may not always present an optimal learning situation (Becker & Jenny, 2017; Shim & Norman, 2015). In all, the current study is a first examination of the efficacy of BeeLine Reader. The application has some potential but the hypotheses that BeeLine Reader makes reading 'easier, faster, and more enjoyable' require further and more broadly-oriented empirical support. Furthermore, the bold claim of the developers of BeeLine Reader that the application 'is a research-backed tool that improves reading ability for students of all ages and skill levels' (BeeLine Reader, 2021) is not supported by the data in our study.

## Credit author statement

**Arnout Koornneef:** conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing - original draft & review & editing, visualization, supervision, project administration. **Astrid Kraal:** investigation, resources, writing - review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[9] Only about 50% of BeeLine Reader's users prefer the default settings (N. Lum, personal communication, June 7, 2021).

## References


Becker, A., & Jenny, S. E. (2017). No need for training wheels: Ideas for including balance bikes in elementary physical education. *Journal of Physical Education, Recreation and Dance, 88*(4), 14–21. https://doi.org/10.1080/07303084.2016.1270789

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

Bernard, J. B., Anne-Catherine, S., & Eric, C. (2007). Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision Research, 47* (28), 3447–3459. https://doi.org/10.1016/j.visres.2007.10.005

Blackmore-Wright, S., Georgeson, M. A., & Anderson, S. J. (2013). Enhanced text spacing improves reading performance in individuals with macular disease. *PLoS One, 8*(11), Article e80325. https://doi.org/10.1371/journal.pone.0080325

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science, 23*(3), 201–207. https://doi.org/10.1177/0963721414530145

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford library of psychology. The Oxford handbook of eye movements* (pp. 643–662). Oxford University Press.

van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language, 27*(1), 1–22. https://doi.org/10.1016/0749-596X(88)90045-9

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5/6), 489–503. https://doi.org/10.1023/A:1008084120205

Calabrèse, A., Bernard, J. B., Hoffart, L., Faure, G., Barouch, F., Conrath, J., et al. (2010). Small effect of interline spacing on maximal reading speed in low-vision patients with central field loss irrespective of scotoma size. *Investigative Ophthalmology & Visual Science, 51*(2), 1247–1254. https://doi.org/10.1167/iovs.09-3682

Chung, S. T. L. (2004). Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optometry and Vision Science, 81*(7), 525–535. https://doi.org/10.1097/00006324-200407000-00014

Chung, S. T. L., Jarvis, S. H., Woo, S. Y., Hanson, K., & Jose, R. T. (2008). Reading speed does not benefit from increased line spacing in AMD patients. *Optometry and Vision Science, 85*(9), 827–833. https://doi.org/10.1097/OPX.0b013e31818527ea

Drummond, A. (2018). *Ibex farm. Online Server: Http://Spellout. Net/Ibexfarm*.

Dyson, M. C. (2004). How physical text layout affects reading from screen, 6. In *Behaviour and information technology* (Vol. 23, pp. 377–393). Taylor & Francis Group. https://doi.org/10.1080/01449290410001715714.

Dyson, M. C., & Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies, 54*(4), 585–612. https://doi.org/10.1006/ijhc.2001.0458

Evers-Vermeul, J. (2020). Short sentences, easy to read? Effects of connectives and layout on text comprehension by beginning readers. In N. Gagarina, & R. Musan (Eds.), *Referential and relational discourse coherence in adults and children* (pp. 41–56). De Gruyter Mouton.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371–395. https://doi.org/10.1037/0033-295x.101.3.371

Henderson, L. M., Tsogka, N., & Snowling, M. J. (2013). Questioning the benefits that coloured overlays can have for reading in students with and without dyslexia. *Journal of Research in Special Educational Needs, 13*(1), 57–65. https://doi.org/10.1111/j.1471-3802.2012.01237.x

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49*(24), 2948–2959. https://doi.org/10.1016/j.visres.2009.09.012

Humar, I., Gradišar, M., & Turk, T. (2008). The impact of color combinations on the legibility of a Web page text presented on CRT displays. *International Journal of Industrial Ergonomics, 38*(11–12), 885–899. https://doi.org/10.1016/j.ergon.2008.03.004

Humar, I., Gradisar, M., Turk, T., & Erjavec, J. (2014). The impact of color combinations on the legibility of text presented on LCDs. *Applied Ergonomics, 45*(6), 1510–1517. https://doi.org/10.1016/j.apergo.2014.04.013

Iovino, I., Fletcher, J. M., Breitmeyer, B. G., & Foorman, B. R. (1998). Colored overlays for visual perceptual deficits in children with reading disability and attention deficit/hyperactivity disorder: Are they differentially effective? *Journal of Clinical and Experimental Neuropsychology, 20*(6), 791–806. https://doi.org/10.1076/jcen.20.6.791.1113

Jeon, S. T., Hamid, J., Maurer, D., & Lewis, T. L. (2010). Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *Journal of Experimental Child Psychology, 107*(4), 423–437. https://doi.org/10.1016/j.jecp.2010.05.009

Joo, S. J., White, A. L., Strodtman, D. J., & Yeatman, J. D. (2018). Optimizing text for an individual's visual system: The contribution of visual crowding to reading difficulties. *Cortex, 103*, 291–301. https://doi.org/10.1016/j.cortex.2018.03.013

Katzir, T., Hershko, S., & Halamish, V. (2013). The effect of font size on reading comprehension on second and Fifth grade children: Bigger is not always better. *PLoS One, 8*(9), Article e74061. https://doi.org/10.1371/journal.pone.0074061

Kim, Y. S. G., Petscher, Y., & Vorstius, C. (2019). Unpacking eye movements during oral and silent reading and their relations to reading proficiency in beginning readers. *Contemporary Educational Psychology, 58*, 102–120. https://doi.org/10.1016/j.cedpsych.2019.03.002

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

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

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come. *Organizational Research Methods, 15*(4), 722–752. https://doi.org/10.1177/1094428112457829

Land, J. (2009). *Zwakke lezers, sterke teksten? : Effecten van tekst- en lezerskenmerken op het tekstbegrip en de tekstwaardering van vmbo-leerlingen*. Eburon. https://dspace.library.uu.nl/handle/1874/32240

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

Ling, J., & van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of web pages. *Displays, 28*(2), 60–67. https://doi.org/10.1016/j.displa.2007.04.003

Madhavan, I., Izzuddin Hairol, M., & Mohammed, Z. (2016). Spacing improves reading in dyslexic children. *Asia Pacific Journal of Developmental Differences, 3*(1). https://doi.org/10.3850/S2345734114000183

Nanavati, A. A., & Bias, R. G. (2005). Optimal line length in reading-A literature review. *Visible Language, 39*(2), 120.

Netchine, S., Guihou, M., Greenbaum, C., & Englander, G. (1983). Retour à la ligne, âge des lecteurs et accessibilité au texte. *Travail Humain, Le, 46*(1), 139–153. http://search.proquest.com/openview/faa31fcd683133c3ef7283510c44a633/1?pq-origsite=gscholar&cbl=1817992.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas—Part II. *Linguistics and Language Compass, 10*(11). https://doi.org/10.1111/lnc3.12207

Parker, A. J., Slattery, T. J., & Kirkby, J. A. (2019). Return-sweep saccades during reading in adults and children. *Vision Research, 155*, 35–43. https://doi.org/10.1016/j.visres.2018.12.007

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135. https://doi.org/10.1038/nn.2187

Petrov, Y., & Meleshkevich, O. (2011). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research, 51*(10), 1117–1123. https://doi.org/10.1016/j.visres.2011.03.001

Pinna, B., & Deiana, K. (2018). On the role of color in reading and comprehension tasks in dyslexic children and adults. *I-Perception, 9*(3). https://doi.org/10.1177/2041669518779098, 204166951877909.

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., & Pollatsek, A. (1989). *Psychology of reading*. Lawrence Erlbaum Associates.

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

Rello, L., Pielot, M., & Marcos, M.-C. (2016). Make it big! the effect of font size and line spacing on online readability. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 3637–3648).

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

Rodrigue, T. K. (2017). The digital reader, the alphabetic writer, and the space between: A study in digital reading and source-based writing. *Computers and Composition, 46*, 4–20. https://doi.org/10.1016/J.COMPCOM.2017.09.005

Schneps, M. (2015). *Using technology to break the speed barrier of reading*. Scientific American MIND. https://www.scientificamerican.com/article/using-technology-to-break-the-speed-barrier-of-reading/.

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

Schneps, M. H., Thomson, J. M., Chen, C., Sonnert, G., & Pomplun, M. (2013). E-readers are more effective than paper for some with dyslexia. *PLoS One, 8*(9), Article e75634. https://doi.org/10.1371/journal.pone.0075634

Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PLoS One, 8*(8), Article e71161. https://doi.org/10.1371/journal.pone.0071161

Schroeder, S., Hyönä, J., & Liversedge, S. P. (2015). Developmental eye-tracking research in reading: Introduction to the special issue, 5. In *Journal of cognitive psychology* (Vol. 27, pp. 500–510). Psychology Press Ltd. https://doi.org/10.1080/20445911.2015.1046877.

Shim, A. L., & Norman, S. (2015). Incorporating pedal-less bicycles into a pre-K through third-grade curriculum to improve stability in children. *Journal of Physical Education, Recreation and Dance, 86*(1), 50–51. https://doi.org/10.1080/07303084.2015.978681

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

Toste, J. R., Didion, L., Peng, P., Filderman, M. J., & McClelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for K–12 students. *Review of Educational Research, 90*(3), 420–456. https://doi.org/10.3102/0034654320919352

Unrau, N. J., Rueda, R., Son, E., Polanin, J. R., Lundeen, R. J., & Muraszewski, A. K. (2018). Can reading self-efficacy Be modified? A meta-analysis of the impact of interventions on reading self-efficacy. *Review of Educational Research, 88*(2), 167–204. https://doi.org/10.3102/0034654317743199

van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. M. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology, 106*(4), 1036–1048. https://doi.org/10.1037/a0036293

van Silfhout, G., Evers-Vermeul, J., & Sanders, T. J. M. (2014). Establishing coherence in schoolbook texts: How connectives and layout affect students' text comprehension. *Dutch Journal of Applied Linguistics, 3*(1), 1–29. https://doi.org/10.1075/dujal.3.1.01sil

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 (computer science working papers)*. Department of Computer Science, University of Waikato. https://hdl.handle.net/10289/802.

Walker, S., Black, A., Bessemans, A., Bormans, K., Renckens, M., & Barratt, M. (2018). Chapter 2. Designing digital texts for beginner readers. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 31–56). John Benjamins. https://doi.org/10.1075/swll.17.02wal.

Williams, M. C., LeCluyse, K., & Rock-Faucheux, A. (1992). Effective interventions for reading disability. *Journal of the American Optometric Association, 63*, 411–417.

BeeLine Reader: making reading on-screen easier and faster, (2021–. (Accessed 18 May 2021).